Manfred J. Holler
Hannu Nurmi *Editors*

Power, Voting, and Voting Power: 30 Years After

Springer

Power, Voting, and Voting Power: 30 Years After

Manfred J. Holler · Hannu Nurmi
Editors

Power, Voting, and Voting Power: 30 Years After

Springer

Editors
Manfred J. Holler
Department of Economics
Institute of SocioEconomics
University of Hamburg
Hamburg, Germany

Hannu Nurmi
Department of Political Science
and Contemporary History
University of Turku
Turku, Finland

ISBN 978-3-642-35928-6 ISBN 978-3-642-35929-3 (eBook)
DOI 10.1007/978-3-642-35929-3
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013931276

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword

Political, economic, social life is essentially governed according to the power of agents, be they individuals, institutions, states, countries, etc. As a consequence, it is not surprising that power is a major ingredient of social science. Although this appears today as self-evident, it was not the case some decades ago. In 1938, no less than Bertrand Russell devoted a volume to this topic. I am afraid that this book has been rather neglected.[1] Russell wrote on page 4: *In the course of this book I shall be concerned to prove that the fundamental concept in social science is Power, in the same sense in which Energy is the fundamental concept in physics*.

Standard microeconomic theory which culminates with the beautiful construction of (Walrasian) general equilibrium theory by Kenneth Arrow, Gérard Debreu, and Lionel McKenzie not only neglects power, but, in some sense, negates it. The general equilibrium framework appears as an ideal situation to which society should tend: perfect competition. The best mathematical tool to model perfect competition was introduced by Robert Aumann. It consists in assuming a continuum of agents, so that each agent's influence (on prices) is negligible. Whatever the formalization, either a finite number of agents, an infinite countable set of agents or a continuum, with perfect competition, agents are so-called price takers. However, in the real world, there exist markets where there are only a few agents (at least on one side of the market) and these agents will possess *market power*. In the microeconomic theory Bible (Mas-Colell et al. 1995), *Market Power* is the title of Chap. 12 (there are 23 chapters). Within Aumann's measure-theoretic framework it has been possible to formalize at the same time negligible and powerful agents. There is a fundamental difficulty to mix the finite and the infinite, the continuous and the discrete and, in spite of remarkable works by Benyamin Shitovitz and others in the 1970s, the mainstream microeconomic research has

[1] Bertrand Russell, *Power. A New Social Analysis, George Allen and Unwin*. My attention was called to this book by a paper of Abraham Diskin and Moshe Koppel: *The Measurement of Voting Power as a Special Case of the Measurement of Political Power*, to appear in *Voting Power and Procedures. Essays in Honour of Dan Felsenthal and Moshé Machover* ed. by R. Fara et al., Springer. To my surprise, I discovered Russell's book in my personal library and I must add, to my shame, that the copy was like new.

followed another route, more or less forgetting the general equilibrium approach, a regression to my view.

The ideal of equality has its social choice theoretic version as *anonymity*. Equality here means basically equality of power. Having equal power for agents does not mean that they have no power (unless we consider that they are elements of a continuum). Rather it means that they have the same power, possibly weak depending on their number. Having the same power leads to the possibility that some agents may have more power than others. Then rather than viewing power as an absolute concept we can consider that it is a relative concept where the different power of various agents can be compared. In the three famous impossibility results of social choice theory (Arrow, Sen and Gibbard–Pattanaik–Satterthwaite Theorems), the notion of power is implicit, hidden in admissible or repulsive concepts. In Arrow's Theorem, given independence of irrelevant alternatives, a sufficient heterogeneity of agents' preferences and some level of collective rationality, there is a consistency problem between unanimity—the fact that all the agents taken together as a group are powerful over all social states—and the absence of dictatorship—the fact that a given single individual is powerful over all these social states. In Sen's theorem (the impossibility of a Paretian liberal) there is an inconsistency between unanimity again and the fact that at least two agents are powerful over at least two social states, this fact being justified by an interpretation in terms of individual rights or freedom of choice within a personal sphere, an idea going back to John Stuart Mill. In Gibbard–Pattanaik–Satterthwaite's theorem, the conflict is basically between non-dictatorship again and the possibility for an agent to obtain a benefit from acting strategically by misrepresenting her 'sincere' preference.

It is certainly in the part of social choice devoted to voting and in (cooperative) game theory that the notion of power has, at last, reached preeminence. Although equality is an ideal in some configurations, it is not in others. This is particularly true when voters represent institutions such as constituencies of different size, states in a federal system, countries, etc. In 1986, William Riker called our attention to Luther Martin (!), a delegate from Maryland to the Constitutional Convention in Philadelphia in 1787. Luther Martin made calculations of the voting power of the (then) 13 American states on the basis of a fictitious weighted voting game in which representatives of a given state voted together. According to William Riker the method he proposed is very similar to what John Banzhaf proposed in the 1960s—or, according to Philip Straffin, to what J. Deegan and E. W. Packel proposed in 1978 or, according to Dan Felsenthal and Moshé Machover, to what Manfred Holler proposed in 1982. (We now know, principally thanks to Felsenthal and Machover's book, that Banzhaf was also preceded by Lionel Penrose). Lloyd Shapley and Martin Shubik developed in the 1950s a game-theoretic approach to the measurement of (voting) power based on the so-called Shapley value. In cooperative game theory, a basic structure introduced by John von Neumann and Oskar Morgenstern is the simple game structure. Groups of agents/players (coalitions) are either powerful or without power. A simple game basically amounts to identify the coalitions which are powerful (called 'winning').

In the real world, these winning coalitions can be established on the basis of a quite strong inequality among the players as within the Security Council of UN where some countries have a veto, the so-called permanent members of the Council (a winning coalition must include all permanent members plus a sufficient fraction of non permanent members who are elected by the General Assembly—the number of non permanent members and consequently the minimum number of non permanent members to form a winning coalition has varied since the creation of UN and the treatment of abstentions of permanent members has been rather ambiguous). Winning coalitions can also be established on the basis of weights given to players when the players are states, countries etc. A remarkable and recent example of the difficulties related to a priori voting power is the choice of weights and quota for the countries in the Council of Ministers of the European Union.

This book is a major contribution to the advancement of our knowledge on power and specifically voting power by some of the most important scholars in this area. The two editors themselves made brilliant contributions to the measurement of power (Manfred Holler has his name associated to a well-known power index to which I previously alluded) and more generally to voting analysis (Hannu Nurmi published a number of books which became classical).

Reference

Mas-Colell, A., Whinston, M. D. & Green, J. R. (1995). *Microeconomic Theory*. Oxford: Oxford University Press

Caen, December 3, 2012

Maurice Salles
maurice.salles@unicaen.fr

Contents

Contributors

Marlies Ahlert Department of Law and Economics, Martin-Luther-University Halle-Wittenberg, Halle, Germany, e-mail: marlies.ahlert@wiwi.uni-halle.de

J. M. Alonso-Meijide Department of Statistics and Operations Research, University of Santiago de Compostela, Santiago de Compostela, Spain, e-mail: josemaria.alonso@usc.es

Rafel Amer Department of Applied Mathematics II and Industrial and Aeronautical Engineering School of Terrassa, Technical University of Catalonia, Colom 11, 08222 Terrassa, Spain, e-mail: rafel.amer@upc.edu

Cesarino Bertini Department of Management, Economics and Quantitative Methods, University of Bergamo, via dei Caniani 2, 24127 Bergamo, Italy, e-mail: cbertini@unibg.it

Matthew Braham Institute of Philosophy, University of Bayreuth, Universitätsstr. 30, 95447 Bayreuth, Germany, e-mail: matthew.braham@uni-bayreuth.de

Thomas Bräuninger University of Mannheim, 68131 Mannheim, Germany, e-mail: thomas.braeuninger@uni-mannheim.de

Francesc Carreras Department of Applied Mathematics II and Industrial and Aeronautical Engineering School of Terrassa, Technical University of Catalonia, Colom 11, 08222 Terrassa, Spain; ETSEIAT, P.O. Box 577, Terrassa 08220, Spain, e-mail: francesc.carreras@upc.edu

Ian Carter Department of Political and Social Studies, University of Pavia, Pavia, Italy, e-mail: icarter@unipv.it

Balbina Casas-Méndez Department of Statistics and Operations Research, University of Santiago de Compostela, Santiago de Compostela, Spain, e-mail: balbina.casas.mendez@usc.es

Vincent C. H. Chua SIM University, Clementi, Singapore, e-mail: vincentchua @unisim.edu.sg

Daniel Eckert Institute of Public Economics, University of Graz, 8010 Graz, Austria, e-mail: daniel.eckert@uni-graz.at

Dan S. Felsenthal University of Haifa, Haifa, Israel; Centre for the Philosophy of the Natural and Social Sciences, London School of Economics and Political Science, Lakatos Building, Houghton Street, London WC2A 2AE, UK, e-mail: dfelsenthal@poli.haifa.ac.il

M. G. Fiestras-Janeiro Department of Statistics and Operations Research, University of Vigo, Vigo, Spain, e-mail: fiestras@uvigo.es

Vito Fragnelli Department of Science and Technological Innovation, University of Eastern Piedmont, Viale T. Michel 11, 15121 Alessandria, Italy, e-mail: vito.fragnelli@mfn.unipmn.it

Josep Freixas Department of Applied Mathematics III and High Engineering School of Manresa, Technical University of Catalonia, EPSEM, Av. Bases de Manresa, 61-73, 08242 Manresa, Spain, e-mail: josep.freixas@upc.edu

Gianfranco Gambarelli Department of Management, Economics and Quantitative Methods, University of Bergamo, via dei Caniani 2, 24127 Bergamo, Italy, e-mail: gambarex@unibg.it

Manfred J. Holler Institute of SocioEconomics, University of Hamburg, Von-Melle Park 5 Hamburg, 20146, Germany; Public Choice Research Centre, University of Turku, Turku, Finland, e-mail: holler@econ.uni-hamburg.de

Madeleine O. Hosli Department of Political Science, Leiden University, P.O. Box 9555, 2300RB Leiden, The Netherlands, e-mail: hosli@fsw.leidenuniv.nl

Ron Johnston School of Geographical Sciences, University of Bristol, University Road, Bristol, BS8 ISS, UK, e-mail: r.johnston@bristol.ac.uk

Serguei Kaniovski Austrian Institute of Economic Research (WIFO), P.O. Box 91, 1103 Vienna, Austria, e-mail: serguei.kaniovski@wifo.ac.at

Werner Kirsch Fakultät für Mathematik und Informatik, FernUniversität in Hagen, Hagen, Germany, e-mail: werner.kirsch@fernuni-hagen.de

Christian Klamler Institute of Public Economics, University of Graz, 8010 Graz, Austria, e-mail: christian.klamler@uni-graz.at

Hartmut Kliemt Frankfurt School of Finance and Management, Frankfurt, Germany, e-mail: h.kliemt@frankfurt-school.de

Thomas König University of Mannheim, 68131 Mannheim, Germany, e-mail: koenig@uni-mannheim.de

Eerik Lagerspetz Department of Behavioural Sciences and Philosophy, University of Turku, 20014 Turku, Finland, e-mail: eerlag@utu.fi

Annick Laruelle BRiDGE, Foundations of Economic Analysis I, University of the Basque Country (UPV/EHU), Avenida Lehendakari Aguirre, 83, 48015 Bilbao, Spain; IKERBASQUE Basque Foundation of Science, 48011 Bilbao, Spain, e-mail: a.laruelle@ikerbasque.org

Dennis Leech Department of Economics, University of Warwick, Coventry CV4 7AL, Centre for the Philosophy of the Natural and Social Sciences, London School of Economics and Political Science, Lakatos Building, Houghton Street, London WC2A 2AE, UK, e-mail: d.leech@warwick.ac.uk

Robert Leech Imperial College, London, UK, e-mail: r.leech@imperial. ac.uk

Peter Leutgäb Institute for Mathematics, University of Augsburg, 86135 Augsburg, Germany, e-mail: Peter.Leutgaeb@GMX.Net

Moshé Machover King's College, London, UK; Centre for the Philosophy of the Natural and Social Sciences, London School of Economics and Political Science, Lakatos Building, Houghton Street, London WC2A 2AE, UK, e-mail: moshe.machover@kcl.ac.uk

Alexander Mayer Department of Economics, University of Bayreuth, 95440 Bayreuth, Germany, e-mail: alexander.mayer@uni-bayreuth.de

Tommi Meskanen Department of Mathematics and Statistics, University of Turku, 20014 Turku, Finland, e-mail: tommes@utu.fi

Nicholas R. Miller Department of Political Science, University of Maryland Baltimore County, Baltimore, MD 21250, USA, e-mail: nmiller@umbc.edu

Maria Montero School of Economics, University of Nottingham, University Park, Nottingham NG7 2RD, UK, e-mail: maria.montero@nottingham.ac.uk

Stefan Napel Department of Economics, University of Bayreuth, 95440 Bayreuth, Germany; Public Choice Research Centre, University of Turku, Turku, Finland, e-mail: stefan.napel@uni-bayreuth.de

Andreas Nohn Public Choice Research Centre, University of Turku, Turku, Finland; Institute of Socio Economics, University of Hamburg, 20146 Hamburg, Germany, e-mail: nohn@econ.uni-hamburg.de

Hannu Nurmi Department of Political Science and Contemporary History and Public Choice Research Centre, University of Turku, 20014 Turku, Finland, e-mail: hnurmi@utu.fi

Rie Ono-Yoshida Faculty of Law and Economics, Chiba University, 1-33 Yayoicho Inage-ku, Chiba 263-8522, Japan, e-mail: onorie@le.chiba-u.ac.jp

Guillermo Owen Department of Mathematics, Naval Postgraduate School, Monterey, California, e-mail: gowen@nps.edu

Antti Pajala Public Choice Research Centre, University of Turku, 20014 Turku, Finland, e-mail: anpaja@utu.fi

Arsen Palestini MEMOTEF, Sapienza University of Rome, Rome, Italy, e-mail: arsen.palestini@uniroma1.it

Montserrat Pons Department of Applied Mathematics III and High Engineering School of Manresa, Technical University of Catalonia, EPSEM, Av. Bases de Manresa, 61-73, 08242 Manresa, Spain, e-mail: montserrat.pons@upc.edu

Friedrich Pukelsheim Institute for Mathematics, University of Augsburg, 86135 Augsburg, Germany, e-mail: Pukelsheim@Math.Uni-Augsburg.De

Donald G. Saari Institute for Mathematical Behavioral Science, University of California, Irvine, CA 92697-5100, USA, e-mail: dsaari@uci.edu

Maurice Salles CREM (UMR - CNRS 6211) and Institute for SCW, University of Caen, Caen Cedex 14032, France; CPNSS London School of Economics, Houghton Street, London WC2A 2AB, UK; Murat Sertel Center for Advanced Economic Studies, Bilgi University, Istanbul, Turkey, e-mail: maurice.salles@unicaen.fr

Evan Schnidman Department of Government, Harvard University, Cambridge, MA, USA, e-mail: evan.schnidman@gmail.com

Norman Schofield Center in Political Economy, Washington University in Saint Louis, 1 Brookings Drive, Saint Louis, MO 63130, USA, e-mail: schofield.norman@gmail.com

Wojciech Słomczyński Institute of Mathematics, Jagiellonian University, ul. Lojasiewicza 6, 30-348 Kraków, Poland, e-mail: wojciech.slomczynski@im.uj.edu.pl

Izabella Stach Faculty of Management, AGH University of Science and Technology, Krakow, Poland, e-mail: istach@zarz.agh.edu.pl

František Turnovec Institute of Economic Studies, Charles University in Prague, Opletalova 26, 110 00 Prague 1, Czech Republic, e-mail: turnovec@fsv.cuni.cz

Federico Valenciano BRiDGE, Applied Economics IV, University of the Basque Country (UPV/EHU), Avenida Lehendakari Aguirre, 83, 48015 Bilbao, Spain, e-mail: federico.valenciano@ehu.es

Hannu Vartiainen HECER, University of Helsinki, Arkadiankatu 7, 00014 Helsinki, Finland, e-mail: hannu.vartiainen@helsinki.fi

Mika Widgrén (deceased) Turku School of Economics and Public Choice Research Centre, University of Turku, 20014 Turku, Finland

Karol Życzkowski Institute of Physics, Jagiellonian University, ul.Reymonta 4, 30-059 Kraków, Poland; Center for Theoretical Physics, Polish Academy of Sciences, Al. Lotników 32/46, 02-668 Warszawa, Poland, e-mail: karol@tatry.if.uj.edu.pl

Reflections on Power, Voting, and Voting Power

Manfred J. Holler and Hannu Nurmi

1 An Introduction to Power, Voting, and Voting Power: 30 Years After

Power is a fundamental concept in the social sciences. It is, however, a theoretical one, i.e., it cannot be directly observed. It is also dispositional. If a person or institution has power, it has an ability or propensity to bring about certain types of events or other outcomes. From a formal point of view, power can be represented as a unary predicate ("A has power") or a binary relation ("A has power over B") or a ternary relation ("A has power over B with regard to X"). Nothing has changed in these fundamental relations since the publication of the volume *Voting, Power and Voting Power* (PVVP) in 1982. But subsequently many articles/studies… derived from the material presented in that volume (PVVP) have been published. Some of those publications directly refer to contributions that can be found in the 1982 collection of chapters. However, this is not the primary argument for publishing a second volume *Voting*, *Power and Voting Power* thirty years later. More convincing to us is that there has been a lot of new material developed during the last thirty years in the fields of PVVP. We think it high time for reflections about what has been accomplished during these years and what are the main issues of ongoing and future research. PVVP 2012 should be of help in answering these questions. Of course, the selection of material is highly subjective.

M. J. Holler (✉)
University of Hamburg, Hamburg, Germany and Public Choice Research Centre, University of Turku, 20014 Turku, Finland
e-mail: holler@econ.uni-hamburg.de

H. Nurmi
Department of Political Science and Contemporary History and Public Choice Research Centre, University of Turku, 20014 Turku, Finland

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*, DOI: 10.1007/978-3-642-35929-3_1,

We find the selected chapters very important contributions. Some are original material written for PVVP 2012. However, most of the chapters are more or less revised material published in the quarterly journal *Homo Oeconomicus*, and are, thus, accessible only for a small readership. We did not select articles that are available in leading and widely distributed journals of social sciences, economics, game theory and mathematics. This of course gives an additional bias to our volume. However, we think that we can leave it to the reader to find the "easy to access" articles in the library—whether in paper or in electronic form.

This is not the only bias that characterizes PVVP 2012. The volume is the result of our ideas about what research work and which results are important (or interesting) and what will be important for the future. This selection, of course, has to do with our own work in this field. However, we will follow, in this introduction and in the selection of the contributions to this volume, the two paths that Anatol Rapoport outlined in his *Foreword* to the 1982 volume: game theory is the one, and social choice theory the other.

Given these two foci and our personal biases, the contributions to this volume reflect the main issues in the discussion of power, voting and voting power over the last thirty years.

2 Power and Preferences

There is an ongoing debate on whether power measures should take the preferences of the agents into account, and if so, to what degree. For instance, the *Journal of Theoretical Politics* (JTP) dedicated many pages of its volume 11 (1999) to this issue. Those authors who wanted to see preferences taken into consideration even declared that power indices are useless—at least when it comes to measuring power in the EU—,[1] while others argued that power indices are valuable instruments just because they do not refer to preferences which might be unknown or irrelevant for the questions under scrutiny. The latter position was defended with reference to institutional design: future agents and their priorities are—or at least often should be—irrelevant and, in any case, unknown for the present-day deliberations. When on March 25th, 1957, the Treaty of Rome was signed creating the European Economic Community (EEC) of *The Six*, and the seats in the Council of Ministers were allocated to the participating countries, the signing partners could not foresee the political preferences of the governments that were to be represented in the coming years. (See Holler and Widgrén 1999a.)

In the course of the scholarly debate that took place, e.g., in JTP, a consensus seemed to emerge suggesting that political preferences are to be considered when power measures are used to forecast or to analyse specific outcomes or events

[1] See Garrett and Tsebelis (1996, 1999), Tsebelis and Garrett (1996, 1997) and Steunenberg et al. (1999) for this message.

defined by specific historical settings,[2] just like other factors that are affecting the outcome. Thus, for example, election outcomes are sometimes interpreted as depending on whether it rained or not. However, despite the apparent consensus the discussion about power and preferences has been popping up time and again. Napel and Widgrén (2005) argue for the "possibility of a preference-based power index," this being the title of their article, while Braham and Holler's (2005a, b) retort is the "impossibility of a preference-based power index".[3]

As Napel and Widgrén use all *possible* single-peaked *preferences*, one could argue that they use the assumed preferences as an analytical device to measure power, and not as an ingredient of power. In fact, it seems that they apply the preference profiles in order to defend their choice of the Shapley-Shubik index which is related to permutations of agents instead of unordered sets of agents, i.e., coalitions. However, even Shapley and Shubik (1954) doubted the *plausibility* of applying the Shapley value to weighted voting. Undoubtedly, information of preferences, whether fully hypothetical or with some empirical substance, can be useful to give us a better understanding of power measures. To put water into a bucket will show us whether the bucket has a hole or not. However, water is not part of the bucket. In many applications we may use the bucket without having it filled with water.

This volume opens with two contributions, the first one authored by Ian Carter (2013) and the second by Matthew Braham (2013), that discuss the nature of power. Conceptual issues are also discussed by Laruelle and Valenciano (2013). For many of the contributions that follow Max Weber's definition of power is a good starting point.[4] Unfortunately, there are somewhat incompatible alternative translations of Max Weber's concept of power. Parsons translated "Macht bedeutet die Chance, …" as "the *probability* that one actor within a social relationship will be in a position to carry out his own will despite resistance." (Weber 1947[1922]: 152, italics added).[5] This is the translation of Weber's definition of power on page 38 of *Wirtschaft und Gesellschaft*, published posthumously (Weber 2005[1922]: 38). In the *Essays from Max Weber*, edited by Gerth and Mills, we read: "In general, we understand by 'power' the chance of a man or of a number of men to realize their own will in a communal action even against the resistance of others who are participating in the action" (Weber 1948[1924]:180). This is the translation of Weber's definition given on page 678 of *Wirtschaft und Gesellschaft*

[2] See, e.g., Aleskerov et al. (2013), analyzing the power distributions in the Weimar Reichstag in 1919–1933.

[3] This debate had a locus: the Institute of SocioEconomics on the second floor of Von Melle Park 5 of the University of Hamburg where three of the four contributors to this debate had their offices and the fourth, Mika Widgrén, was a regular visitor.

[4] This section on Weber derives from Holler and Nurmi (2010).

[5] The German text is "Macht bedeutet jede Chance, innerhalb einer sozialen Beziehung den eigenen Willen auch gegen Widerstreben durchzusetzen, gleichviel worauf diese Chance beruht"(Weber 2005[1922]: 38).

(Weber 2005[1922]: 678).[6] There are also differences in the two definitions in their original German versions. For instance, the second definition extends the definition of power to "a number of men" and "communal action." Another obvious difference is in the translation of the German word "die Chance" (which, of course, the Germans borrowed from French). Parsons used "probability" for its translation into English while in the edition of Gerth and Mills we read "chance." Quite similar to English, in German "die Chance" expresses either a possibility or is a synonym for probability. It depends on the context whether the former or later interpretation applies. This also holds in the case of Weber's definition of power and the use of "die Chance" in it.

There is a widely shared notion of probability which relates this concept to a random mechanism as, for example, in the expression "chance setup." The outcome of the setup or mechanism is determined by "nature." Chance presupposes a lack of control due, e.g., to decisions or actions of others or to unpredictable natural events. However, if somebody asks "what is the chance to see you tomorrow," an answer "with probability 1/3" does not make sense if the answer solely depends on your choice. However, it would make perfect sense if you cannot leave the house if it rains and the probability of rain is 2/3.

Experts on Weber claim that his use of "die Chance" concurs with possibility or potential. On the other hand, the fact that Parsons used "probability" for the translation of "die Chance" cannot be neglected. Talcott Parsons received a doctorate from the University of Heidelberg in 1927. The title of his doctoral dissertation was "'Capitalism' in recent German literature: Sombart and Weber."[7] In this volume, we find both interpretations. The idea of power as a potential was emphasized in Holler and Widgrén (1999b) where the value of the characteristic function in a coalitional game is interpreted as power. (See also Napel et al. 2013)

3 The Right Index

The ambiguity in the interpretation of power carries over to the question of the "right index." Over many pages and years, the question of right index focused on a comparison of or, should we say, a competition between, the Shapley-Shubik index and the Banzhaf index—the latter also labelled as Penrose-Banzhaf or Banzhaf-Coleman index—ignoring other candidates like the measures suggested by Johnston (1978), Deegan and Packel (1979) and Holler (1982c). The Banzhaf faction was spearheaded by Dan Felsenthal and Moshé Machover while the Shapley-Shubik index had, e.g., Stefan Napel and the late Mika Widgrén as

[6] "Unter'Macht' wollen wir dabei hier ganz allgemein die Chance eines Menschen oder einer Mehrzahl solcher verstehen, den eigenen Willen in einem Gemeinschaftshandeln auch gegen den Widerstand anderer daran Beteiligten durchzusetzen" (Weber 2005[1922]: 678).

[7] See Parsons (1928, 1929) for its publication in *The Journal of Political Economy.*

eminent supporters despite the fact that the latter two introduced themselves a power measure based on the "inferior player concept" (Napel and Widgrén 2001, 2013). On the other hand, Laruelle and Valenciano gave a new axiomatization for the two indices with axioms that "are remarkably close" such that "both indices appear on the same footing when they are interpreted as measures of power in collective decision-making procedures" (Laruelle and Valenciano 2001: 103). However, as Aumann (1977: 471) observes: "…axiomatics underscores the fact that a 'perfect' solution concept is an unattainable goal, a *fata morgana*; there is something 'wrong', some quirk with every one." Still, axiomatizations "serve a number of useful purposes. First, like any other alternative characterization, they shed additional light on a concept and enable us to 'understand' it better. Second, they underscore and clarify important similarities between concepts, as well as differences between them."

Felsenthal et al. (1998) and Felsenthal and Machover (1998) suggested a compromise, but also differentiation, through the claim that the Banzhaf index describes I-Power, an agent's potential influence over the outcome, whereas the Shapley-Shubik index represents P-Power, an agent's expected share in a fixed prize. However, Turnovec (2004) demonstrated that the distinction does not hold: both measures can be interpreted as expressing I-Power or, alternatively, P-Power. Indeed, these measures can be modeled as values of cooperative games and as probabilities of being 'decisive' without reference to game theory at all. The basic point being that 'pivots' (Shapley-Shubik index) and 'swings' (Banzhaf index) can be taken as special cases of a more general concept of 'decisiveness' (see Turnovec et al. 2008; see also Laruelle and Valenciano 2013 and König and Bräuninger 2013).

Still, the distinction of I-Power and P-Power contributes to the discussion of power measures and often serves as a valuable instrument to structure our intuition. Yet, in the light of Turnovec's results, it is perhaps not a major flaw for the Public Goods Index (PGI), introduced by Holler (1982c, 1984), that Felsenthal and Machover (1998) classify it among the P-Power measures. From Paul Samuelson we learn that there is nothing to share in the case of pure public goods. It is difficult to see why the PGI does not qualify as an I-Power measure like the Banzhaf index does. Loosely speaking, the difference between the PGI and the normalized Banzhaf index boils down to those winning coalitions that are not minimal. Holler (1982c, 1998) argues that these coalitions should not be considered because they imply a potential to freeride if the decisions concern public goods—as is often the case in policy making.[8] This does not mean that surplus coalitions do not form, but they should not be considered when measuring power.

[8] The *basic principles* underlying the public good index are (a) the public good property, i.e. nonrivalry in consumption and nonexcludability of access, and (b) the nonfreeriding property. It is obvious from these principles that (strict) minimum winning coalitions should be considered when it comes to measuring power. All other coalitions are either non-winning or contain at least one member that does not contribute to winning. If coalitions of the second type form, then it is by *luck* or because of similarity of preferences, tradition, etc.—*but not because of power*, as there is a potential for freeriding. (See Holler 1998.)

There is another, more critical remark in Felsenthal and Machover that relates to the PGI. They argue that any a priori measure of power that violates local monotonicity is 'pathological' and should be disqualified from serving as a valid yardstick for measuring power (Felsenthal and Machover 1998: 221ff)—and they correctly point out that the PGI and the Deegan-Packel index violate this property. Holler and Napel (2004a, b) hypothesize that the PGI exhibits nonmonotonicity (and thus confirms that the measure does not satisfy local monotonicity) if the game is not decisive, as the weighted voting game $v° = (51; 35, 20, 15, 15, 15)$ with a PGI of $h° = (4/13, 2/13, 3/13, 3/13, 3/13)$ demonstrates, or is improper and therefore indicates that perhaps we should worry about the design of the decision situation.[9] The more popular power measures, i.e., the Shapley-Shubik index and the Banzhaf one, satisfy local monotonicity and thus do not exhibit any peculiarities if the game is not decisive or is improper. To what extent the PGI can serve as an indicator, revealing certain peculiarities of a game, has been discussed in Holler and Nurmi (2012b).

Interestingly, the Shapley-Shubik and Banzhaf index also violate local monotonicity if we consider a priori unions and the equal probability of permutations and coalitions, respectively, no longer applies.[10] The concept of a priori unions or pre-coalitions is rather crude because it implies that certain coalitions will not form at all, i.e., they have a zero probability of forming. Note since the PGI considers minimum winning coalitions (MWCs) only, this is formally equivalent to putting a zero weight on coalitions that have surplus players. Is this the ("technical") reason why the PGI may show nonmonotonicity?

Instead of accepting the violation of monotonicity, we may ask under which circumstances or decision situations the PGI guarantees monotonic results—this may help to design adequate voting bodies. In Holler et al. (2001), the authors analyze alternative constraints on the number of players and other properties of the decision situations. For example, it is obvious that local monotonicity will not be violated by any of the known power measures, including PGI, if there are n voters and n-2 of these are dummies. It is, however, less obvious that local monotonicity is also satisfied for the PGI if one constrains the set of games so that there are only $n-4$ dummies. A hypothesis that needs further research is that the PGI does not show nonmonotonicity if the voting game is decisive and proper and the number of decision makers is smaller than 6.[11]

Which index is the right one? Many contributions to this volume shed light on this question, e.g. Felsenthal and Machover (2013a, b); Laruelle and Valenciano (2013); König and Bräuninger (2013); Alonso-Meijide et al. (2013a, b); Amer and

[9] For further discussion, see Sect. 8 below.

[10] See Alonso-Meijide and Bowles (2005) for examples of voting games with a priori unions that violate local monotonicity.

[11] Perhaps this result also holds for a larger number of decision makers but we do not know of any proof. For a related discussion and the introduction of *weighted monotonicity*, see Alonso-Meijide and Holler (2009).

Carreras (2013); Widgrén and Napel (2013); Montero (2013) and Freixas and Pons (2013). A possible answer is due to Aumann (1977: 464): "None of them; they are all indicators, not predictors. Different solution concepts are like different indicators of an economy; different methods for calculating a price index; different maps (road, topo, political, geologic, etc., not to speak of scale, projection, etc.); different stock indices (Dow Jones)… They depict or illuminate the situation from different angles; each one stresses certain aspects at the expense of others." We subscribe to this perspective. "Different solution concepts can…be thought of as results of choosing not only which properties one likes, but also which examples one wishes to avoid" (Aumann 1977: 471).

4 Cooperative Games, Bargaining Models and Optimal Strategies

Power indices can be distinguished by their underlying assumptions on coalition formation as well as by the weights they give to these coalitions. The weights may reflect the probabilities that particular coalitions form. Inasmuch as these measures are exogenously given by the rules implicit in the power measure we are in the realm of cooperative game theory. Recently, series of chapters have been published taking into account a priori unions, building on the pioneering articles of Owen (1977, 1982). See the contributions of Alonso-Meijide et al. (2013a, b).

Once we ask the question of whether coalition A forms and why coalition B does not, we enter the domain of noncooperative game theory. A lot of work has been done to derive the standard power indices from bargaining games or to interpret solution concepts that are based on notions of bargaining as power indices—also in order to understand the problem of implementing a given (possibly "fair") power distribution. Maria Montero's (2013) contribution to this volume, proposing the nucleolus as a power index, is an example of the latter. Another chapter by her that deals with the "noncooperative foundations of the nucleolus in majority games" (Montero 2006) obviously represents this same approach. Same is true of Yan's (2002) modelling of a "noncooperative selection of the core," while the contributions by Andreas Nohn (2013) as well as by Francesc Carreras and Guillermo Owen (2013) are examples that fall in the first category. The search for a noncooperative foundations of bargaining power and its relationship to the Shapley-Shubik index in Laruelle and Valenciano (2008) as well as the bidding models in Pérez-Castrillo and Wettstein (2001) and, with some reservation, in Hart and Mas-Colell (1996), supporting the Shapley value, also fall in this category.

The main result in Nohn (2013) is that veto players either hold all of the overall power of 1, or hold no power at all. This somehow reflects the preventive power measure ("power to block") suggested in Coleman (1971). However, power indices also deal with the power to initiate and therefore will, in general, not allocate all the power to veto players. The difference is that in Coleman as well as in any other classical power measure the focus is on winning coalitions, i.e., sets of agents that have the means to accomplish something. To be potentially a member of such a coalition represents the chance that the corresponding agent "within a social relationship will be in a position to carry out his own will despite resistance," borrowing from Weber's definition given above, and thus power. In bargaining models the forming of coalitions is a possible, but not a necessary result. Of course, by definition, a veto player has the potential to block any winning coalition, but other arrangements may also lead to a break down of bargaining and to a zero outcome, which represents zero power. Veto power is important in bargaining games because the standard requirement for agreement is unanimity, but in general not all players are active all the time.

The n-person bargaining models in the tradition of Rubinstein or Baron and Ferejohn do not consider binding coalitions as the point of departure, but the power indices do so. No wonder that bargaining models and power measures are difficult to reconcile. Without being more explicit about coalition formation, the bargaining models are not likely to be successful in giving a noncooperative underpinning to the power indices.

Similar problems are relevant for those approaches that do not apply power measures to express a priori (voting) power but model the interaction of the agents as a game and look for possible equilibria. They substitute the potential of a coalition by a game form and preferences that allow specifying a Nash equilibrium (or a refinement of it) that describes the allocation of payoffs and thereby specifies the power of the players in this game. The analysis of EU codecision-making in Napel et al. (2013) is an example of this approach. Of course, the results depend on the assumed payoff functions. But whether we can generalize the outcome also depends on the structure of decision-making and on the information that the voters have. The assumption that the policy space is one where the voters have single-peaked preferences, face only binary agendas and are endowed with complete information is convenient but hardly descriptive of real world voting bodies. A rather extensive literature shows that, for a given preference profile, the voting outcome may strictly depend on whether we apply plurality voting, Borda count, amendment voting, approval voting or some other voting procedure. Moreover, a slight perturbation of the preferences may change the winning platform and thus the winning coalition to their opposite. (See, e.g., Holler and Nurmi 2012a, b) It has been said that power index analysis hardly ever deals with more complex voting rules and the information of the agents. But at least it does not suffer from the vulnerability to perturbation of preferences as long as the working of the rules does not depend on particular properties that the preferences have to satisfy so that we get a voting outcome at all. The contributions on the aggregation of preferences (Part VI) to this volume clarify some of these problems. We will come back to this issue below.

5 Fair Representation and Mechanism Design

Despite—or perhaps because of—the multitude of indices, on the one hand, and the implementation problems that we just outlined, on the other, the issue of fair representation has been much discussed during the last three decades. One reason is the emergence of and important advances in the field of theoretical mechanism design (see also Saari (2013) and Vartiainen (2013)).[12] Another is the ongoing discussion of adequate institutions for international organizations, such as the European Union (see König and Bräuninger (2013) as well as the contributions in Part V in this volume), the European Central Bank, the IMF (see Leech and Leech (2013)) and the World Bank, and various arrangements (frameworks like the UNFCCC) that deal with climate change and environment policy. (See, e.g., Holler and Wegner (2011) for the latter.) A parallel discussion we find in the business world: the issue of an adequate representation of the stakeholders in the various boards of a firm which, in the case of conflict, make use of voting. [See, e.g., Leech (2013); Gambarelli and Owen (1994, 2002).]

However, most vigorous is the discussion in the political arena. In modern democracies, fair representation is, at least, a two-stage problem that relates votes to seats and thus the vote distribution to the power distribution in the representative voting body.[13] One of the central issues addressed has been whether the influence over the outcomes (e.g., legislation) can be distributed precisely according to the resources (e.g., voting weights) when the rules of decision-making are taken into account. In proportional representation systems this issue has been dealt with by aiming at a reasonably close resemblance between the distribution of support for parties and the distribution of the party seats in the legislature. Upon closer inspection, however, the aim at proportionality turns out to be both ambiguous and vague. It is ambiguous in the sense that proportionality may refer to different things. An outcome that is proportional in one sense may not be proportional in another. The aim at proportionality is vague in the sense that—given a precise interpretation of the concept—the outcomes may exhibit different degrees of proportionality. Thus, for example, Jefferson's (d'Hondt's) method of proportional representation tends to be biased towards larger parties when compared with Webster's (Sainte-Laguë).

The ambiguity of proportionality, in turn, can be illustrated by an example that refers to the preference profile in Table 1. Suppose that two candidates out of four (A, B, C and D) are to be elected. If the preferences given above are those reported by the voters, the plurality outcome is {A, B}, whereas proportionality when viewed from the perspective of the Borda count is {C, D}. i.e., depending on the

[12] See the 2007 Nobel Prizes for Leonid Hurwicz, Eric Maskin and Roger Myerson.

[13] In fact, the problem of fair representation can be extended even further because, in general, not everybody is allowed to vote. Minors can be viewed as an instance of such restriction when it comes to voting. Another case is given by felon disenfranchisement. See DeParle (2007: 35) for an illustration and discussion.

Table 1 Preference profile

4,000 voters	3,000 voters	2,000 voters	1,000 voters
A	B	C	A
C	D	D	D
D	C	B	C
B	A	A	B

interpretation of proportionality we may get mutually exclusive choice sets. As proportionality is by and large identified with fair allocation, this result is quite challenging.

Many contributions in *The Logic of Multiparty Systems*, edited by Holler (1987a), analyze the assignment of votes to seats in the case of two or more "criteria of proportionality." In the present volume Gambarelli and Palestini (2013) discuss a multi-district apportionment model that relies on minimax method that in the case of "unavoidable distortions" minimizes the "negative effects." However, voting power is not dealt with in this model.

In the advent of the European Union taking the first steps of enlargement to the Central and Eastern Europe, Laruelle and Widgrén (1998) ask, to paraphrase the title of their chapter, whether the allocation of voting power among EU states is fair. To discuss this question they make use of the Square Root Rule and the Banzhaf index. The relationship between the two will be further discussed in the next section. What is important here is that applying the results of this approach implies a "re-weighting of votes and voting power in the EU," to paraphrase the title of Sutter (2000) that was written as a critical response to Laruelle and Widgrén.

The re-shuffling of seats has been widely discussed in the EU context and, as we will see below, quite a few applications of analytic results have been presented (see Johnston (2013); Kirsch (2013); Bertini et al. (2013); Felsenthal and Machover (2013b) in this volume). History shows that such a policy is accompanied with frustration. Moreover, the re-shuffling method does not always allow perfect proportionality of votes and power. Let us assume a vote distribution $w^\circ = (40, 30, 30)$. Given simple majority voting, there is no re-shuffling of seats so that the corresponding power measure π° is identical with w°, irrespective of whether we apply the Shapley-Shubik index, the Banzhaf index or the PGI. In the introduction to *Power, Voting and Voting Power*, Holler (1982b) gives this example and suggests the randomized decision rule (3/5, 2/5) which prescribes a 3/5 probability of the simple majority and a 2/5 probability for a qualified majority of 2/3 of votes. Here, in order to keep the example simple, the PGI is applied, as it is very easy to list the complete set of minimum wining coalitions for this example. As a result we get the power distributions (1/3, 1/3, 1/3) for the simple majority rule and (1/2, 1/4, 1/4) for the 2/3 quota. Taking care of the randomization (3/5, 2/5) an expected power of $\pi^\circ = (40, 30, 30)$ follows—in percentages, of course.

The randomized decision rule approach was further elaborated in Berg and Holler (1986) and in Holler (1985, 1987b) and generalized in Turnovec (2013).

6 The Case of EU

Unsurprisingly, the issue of fair allocation of seats has been much debated in the context of the enlargement of the EU. In fact the analysis of the EU became the testing field and source of inspiration for almost all questions discussed so far. Therefore we think it appropriate to dedicate more than one page of this introduction to this subject.

6.1 The European Parliament

The enlargement of the EU entitles the new member states to voting rights in the European Parliament (EP) and the Council of Ministers (Council). For the EP, the standard procedure takes into account the size of the population and aims at guaranteeing the representation of the major political parties of each country.[14] Bertini et al. (2013) propose to restructure the distribution of the EP seats according to not only the population sizes but also the economic performance as measured by GDP. They suggest a formula that is based on the Banzhaf index and thus incorporates the potential to form a winning coalition, i.e., a priori voting power. Applying this to the Union of the 27 they show that, with the exception of Italy, all countries have their maximum power value if they either are represented in accordance with their population or, alternatively, with their GDP. The authors do not give a definitive method for allocating seats. Their intention is to build up scenarios to understand which EU country will benefit, if we take into account only GDP, only population, or a linear combination of the two. Taking into account GDP only, the analysis shows that Germany should have 24.35 % of the seats, France 16.38 %, Italy 13.42 %, and so on. This percentage for Italy will decrease if a higher weight is given to the population. It will fall to 12.00 % if only population is taken into consideration. The situation for Poland is quite the opposite: there would be 1.80 % of seats to it if the apportionment is based on GDP, whereas based on population only its share would be 8.04 %.

However, seat shares are notoriously a poor proxy for a priori voting power. Applying the Banzhaf index, Bertini et al. (2013) show that the maximum power for Italy is 12.09 %. This value is not reached in accordance with the maximum number of seats (13.42 %), but through a linear combination $S = 0.8P + 0.2G$ where P and G represent "population" and "GDP," respectively. This linear combination should be Italy's preferred method for assigning seats among EU

[14] Today the EP has 736 members. Of these, 96 members are elected by German voters. The voters of Cyprus, Estonia, Luxembourg and Malta are represented by 6 members each. Since each member state is allocated a much smaller number of seats than in its national parliament, it is inevitable that the smallest parties in each country typically have no representation in the EP, no matter how proportional the election system.

member countries. However, in this case Italy would have only 12.28 % of the seats. For the corresponding voting game its Banzhaf index shows a maximum. (N.B.: all other EU member states can be expected to prefer a different apportionment rule than Italy.)

Here the nonmonotonicity of power is due to the multi-dimensionality of the reference space for the seat apportionment.[15] Individual voters also face the multi-dimensionality of the EP, but in general they are not informed about individual decisions of the EP and the decisions of their representatives. Moreover elections to the EP are often used as by-elections sanctioning the performance of the political parties on the national level.

6.2 The Council of Ministers

The recent history of the shaping of the Council is highlighted by the Nice Treaty of 2001 and the Brussels agreement of 2004. The latter was designed as part of the Treaty establishing a constitution for Europe. The discussion was about the proposed seat distributions, on the one hand, and the decision rules, on the other. In accordance with the Treaty of Nice each EU member state is assigned a voting weight which to some degree reflects its population. With the sum of the weights of all 27 member states being 345, the Council adopts a piece of legislation if following three conditions are satisfied: (a) the sum of the weights of the member states voting in favor is at least 255 (which is approximately a quota of 73.9 %); (b) a simple majority of member states (i.e. at least 14) vote in favor; (c) the member states voting in favor represent at least 62 % of the overall population of the European Union.

The distribution of weights shows, to pick out some prominent features, an equal distribution of 29 votes to the four larger EU member states Germany, France, the UK, and Italy and 4 votes for each of the member states at the opposite end of the scale: Latvia, Slovenia, Estonia, Cyprus and Luxembourg. Malta with a weight of 3 and population of about 400.000 concludes the scale. Note that Germany, with a population of about 82.5 million, and Italy, with a population of 57.7 million, have identical voting weights. The voting weights are monotonic in population size, but obviously this monotonicity is "very" weak.

Condition (c) was meant to correct for imbalances in the ratio of population and seat shares. However, Felsenthal and Machover (2001) demonstrate that the probability of forming a coalition which meets condition (a) but fails to meet one of the other two is extremely low. Therefore, the "triple majority rule" implied by the Nice Treaty boils down to a single rule.

[15] For a more intensive treatment of the multi-dimensionality of the policy space see, e.g., Schofield (2009, 2013).

Given the shortcomings of the voting rule of the Treaty of Nice a revision did not come as a surprise.[16] According to the Brussels agreement of 2004, the Constitutional Treaty, the Council takes its decisions if two criteria are simultaneously satisfied: (a) at least 55 % of EU member states vote in favor; and (b) these member states comprise at least 65 % of the overall population of the EU.

A major defect of the Nice voting rule seems to be the high probability that no decisions will be taken and the status quo prevails, i.e., the decision-making efficiency is low when measured by the Coleman power of a collectivity to act. This measure, the so-called passage probability, represents the probability that the Council would approve a randomly selected issue, where random means "that no EU member knows its stance in advance and each member is equally likely to vote for or against it" (Baldwin and Widgrén 2004: 45). It is specified by the proportion of winning coalitions assuming that all coalitions are equally likely. For the Treaty of Nice rule this measure is 2.1 % only, while for the Constitutional Treaty it is 12.9 %. However, Baldwin and Widgrén (2004) demonstrate that with no substantial change in the voting power of the member states, the Treaty of Nice system can be revised so that its low decision-making efficiency increases significantly. Thus, the difference in effectiveness does not necessarily speak for the Constitutional Treaty rule. But perhaps fairness does.

Condition (b) of the Constitutional Treaty implies that the voting weights applied are directly proportional to the population of the individual member states. At a glance this looks like an acceptable rule, representing the "one man, one vote" principle. However, it caused an outcry in those countries that seem to suffer by the redistribution of a priori voting power implied in the substitution of the "triple majority rule" of the Treaty of Nice by the "double majority rule" of the Constitutional Treaty—also referring to a violation of the "one man, one vote" principle. For instance, Słomczyński and Życzkowski (2007a, b); see also Życzkowski and Słomczyński (2013) in this volume point out that the larger and the smaller countries will gain power should the double majority rule of the Constitutional Treaty prevail, while the medium-sized countries, especially Poland and Spain, will be the losers in comparison to the voting power implications of the Treaty of Nice. (But obviously the Council's voting system of the Treaty of Nice was considered defective.)

Both the Treaty of Nice and the Constitutional Treaty imply voting rules that are based on a compromise between the two principles of equality of member states and equality of citizens. The double majority rule emphasizes these principles. Large states gain from the direct link to population, while small countries would derive disproportionate power from the increase in the number of states needed to support a proposal. The combined effect reduces the a priori voting power of the medium-sized countries. More specifically, Germany will gain by far the most voting power under the Constitutional Treaty rule, giving it 37 % more

[16] Illuminating historical details about the decision-making that led to the Treaty of Nice of 2001 and the Constitutional Treaty of 2004 are described in Baldwin and Widgrén (2004). Obviously, the authors had some inside knowledge.

clout than the UK, while both countries have equal voting power in accordance to rule (a) of the Treaty of Nice. Moreover, the Constitutional Treaty rule will make France the junior partner in the traditional Franco-German alliance which may lead to severe tensions in this relationship.

Obviously, there are substantial differences between the two schemes discussed, and their application to EU decision-making might have substantial and unwarranted consequences. Moreover, there are conflicts of interests made obvious by the analysis of voting power. In order to lessen these conflicts, Słomczyński and Życzkowski (2007a, b, and 2013) propose an allocation of seats and power that they call the "Jagiellonian compromise," named after their home university in Krakow. The core of this compromise is the square root rule, suggested by Penrose (1946). This rule is meant to guarantee that each citizen of each member state has the same power to influence EU decision-making.[17] Applied to the two-tier voting problem of the Council (i.e., voting in the member states at the lower level and in the Council at the upper level), it implies choosing the weights that are proportional to the square root of the population. What remains to be done is to find a quota (i.e., decision rule) such that the voting power of each member state equals its voting weight. But, as already noted, for smaller voting bodies this generally cannot be achieved when applying one quota only. However, the EU has a sufficiently large number of members so that this equality can be duly approximated. Słomczyński and Życzkowski (2007b) give an "optimal quota" of 61.6 % for the EU of 27 member states. Interestingly, the optimal quota decreases with the size of the voting body.[18]

A further expansion of EU membership (e.g., the admission of Turkey) does not constitute a challenge to the square root rule. The adjusted seat distribution will take care of (the square root of) the additional population share, by redistributing seats or by adding additional seats to the Council, and the quota will be revised so that the a priori power is as equal as possible to the seat distribution. This is why Słomczyński and Życzkowski (2007a, b) suggest not fixing the quota in a new constitutional contract but only prescribe a procedure, which assures that (a) the voting weights attributed to each member state are proportional to the square root of the population; and (b) a decision is taken if the sum of the weights of the members that vote yes exceeds the quota $q = 1/2 + 1/\sqrt{\pi M}$, where M represents the number of member states.

The choice of the optimal quota guarantees that the Council's decision-making efficiency of the square root system is always larger than 15.9 %. This is larger

[17] Of course, in all practical terms, this probability is zero. Therefore, the norm of equal power cannot be justified on the basis of potential influence. However, fairness could be a better explanation: individual agents might be powerless, but they do not envy each other.

[18] This is immediate from the approximation of the optimal quota q given in Słomczyński and Życzkowski (2007a). For a voting body of M voters it is: $q = 1/2 + 1/\sqrt{\pi M}$. For the EU of 25 member states the optimal quota was 62 %. (See Słomczyński and Życzkowski (2007b). Also compare Słomczyński and Życzkowski (2006) and Życzkowski and Słomczyński (2013) this volume.).

than calculated for the Constitutional Treaty, and far more than promised by the Treaty of Nice rule. Słomczyński and Życzkowski (2007b) point out that the efficiency of the square root system does not decrease with an increasing number of members states, whereas the efficiency of the double majority rule does.

6.3 Codecision-Making

There is still a puzzle to solve: Why does the allocation of the budget follow national voting power distribution in the Council, as demonstrated by Kauppi and Widgrén (2004, 2007), when the annual spending plans are negotiated between the EP and the Council on the basis of a proposal by the Commission? The EP is organized along ideology based party groups and members of the EP are said not to follow narrowly defined national interests. Is the Council the stronger institution although both institutions are meant to have equal influence on the budget?

Napel and Widgrén (2006), (see also Napel et al. 2013) analyze the power relations of the Council and the EP in the EU legislation under the codecision procedure as a noncooperative game, i.e., both institutions are assumed to act strategically. Their results are that (a) the procedure favors the status quo and (b) the Council has a stronger a priori influence on the outcome than the EP. Both results are due to the qualified majority rule of the Council (whereas the EP only applies simple majority voting). Thus the low decision-making efficiency of the Council, discussed above, carries over to the codecision procedure.

At some stage of the sequential game that the Council and the EP play in the model of Napel and Widgrén, Conciliation Committees enter the arena. Such a committee is composed of the representatives of EU member states—at the time of the study these numbered 25—representing the Council and a delegation of EP members of the same size. It is interesting to note that here the Union of States principle reflected in rule (b) of the Treaty of Nice determines the representation of the Council. This is generally not taken into consideration when the a priori voting power distribution in the Council is analyzed as a weighted voting game. On the other hand, Napel and Widgrén have, in addition to making use of stylized procedural rules that determine the strategies of the players, made some simplifying assumptions on the preferences of the players, i.e., the Council, the EP and the Conciliation Committees, to get a full description of a game model. The individual members of the Council and the EP, also when they are members of a Conciliation Committee, are assumed to have single-peaked preferences. Of course, the latter is a strong assumption, given that many EU policies have a strong distributional character and thus are prone to cyclical majorities and unstable voting outcomes. The fact that we cannot observe a high degree of instability, resulting in prevalent revisions of decisions, seems to be the result of extensive logrolling. The Franco-German alliance is a manifestation of such a policy.

7 Social Choice and Paradoxes of Representation

Voting is a mechanism of aggregating preferences. It also forms a link between the two approaches to voting power singled out by Rapoport in his *Foreword* to PVVP 1982, i.e., game theory and social choice, mentioned above. Voting is often modelled as a game with voters as players and ballots as strategies to choose from. However, voting is a very imperfect way of aggregating preferences if we impose the conditions that Arrow used in his General Possibility Theorem (1963[1951]). Overall, the social choice theory is notorious for its many negative results that demonstrate the incompatibility of various choice desiderata. The outcomes of aggregation under given choice rules do not always seem to reflect the individual opinions in a plausible way. Should we then take preferences into account at all when discussing social decision mechanisms? Although it can be debated whether the analysis of power should take preferences into consideration, it seems obvious that decisions reflect preferences (and perhaps power). At least they should, lest the fundamental democratic principle of "going to the people" be undermined. This should also apply to collective decisions, based on social preferences, unless we argue that "policy is merely a random business." Reasonable choice rules establish a relationship between individual preferences and social preferences, but, as Arrow proved, this relationship is not always straightforward. Social preferences that have the same properties as individual preferences may not exist. In particular, majorities may exhibit properties that would be regarded as irrational when found in individuals. The Condorcet cycle teaches us that the pairwise majority aggregation of individual preference relations, which satisfy transitivity, may lead to intransitivity in the aggregate. While $(A \succ B)$ & $(B \succ C) \Rightarrow (A \succ C)$ is widely accepted as minimum requirement of rational behaviour, and not only by social choice theorists, it could well be that we get a Condorcet cycle $(A \succ B)$ & $(B \succ C) \,\&\, (C \succ A)$ for the society, when aggregating well-ordered individual preferences. We get intransitivity for the social preferences and, as a result, inconsistent decisions. However, to conclude that the society is "irrational" puts too much individualism on it. There are different groups behind the social rankings $(A \succ B)$, $(B \succ C)$, and $(C \succ A)$: $(A \succ B)$, might be supported by a majority that consists of x- and y-voters, $(B \succ C)$ might be supported by a majority that consists of x- and z-voters, and $(C \succ A)$ might be supported by a majority that consists of y- and z-voters, all voters choosing in accordance to their preference order.

Saari (2013) gives a general characterization for preference profiles that will result in such a cycle as just described by the concept of Ranking Wheel Configuration (RWC). Those preference profiles that do not form a RWC are "strongly transitive." The RWC construction provides a way to understand basically all paradoxical results that are related to pairwise or, more generally part-wise, comparisons of alternatives. Eckert and Klamler (2013) apply Saari's geometric approach to discuss paradoxes of majority voting. Ahlert and Kliemt (2013) demonstrate that "numbers may count" (e.g., of victims) in case of the ethical ranking of possible state of affairs. Ono-Yoshida (2013) tests selected solution concepts for multi-choice games and fuzzy games under the assumption that coalitions are binding. This assumption is paired with a bargaining model by Carreras and Owen (2013) who examine the possible proportionality of the Shapley rule, thus matching a model "where only the whole and the individual utilities matter", assuming transferable utilities, with a concept that assumes coalition formation. The subsequent contribution by Vartiainen (2013) does not consider coalitions. However, the log-rolling equilibrium in Vartiainen can be identified with a grand coalition. The failure to achieve such a result implies that the society (i.e., the set of players) remains in the state of anarchy. In the concluding chapter, Schofield (2013) discusses instability and chaos of social decision-making, resuming the coalition framework, and illustrates the implication of the corresponding solution concepts with reference to climate change. This demonstrates a high analytical potential of the social choice tool kit even in the case of anarchy and chaos.

8 Power, Causality, and Responsibility

The concluding section of our reflections deals with an issue which is only indirectly covered by the contributions to this volume, i.e., the allocation of responsibility in collective decision-making.[19] This is motivated by the expectation that if the allocation of responsibility works, threats of punishment or promises of appreciation and honors may improve the results of collective decision-making. However, the specification of causality in the case of collective decision-making with respect to the individual agent cannot be derived from the action and the result as both are determined by the collectivity. They have to be traced back to decision-making itself. But collective decision-making has a quality that differs substantially from individual decision-making. For instance, an agent may support his favored alternative by voting for another alternative or by not voting at all. The two volumes by Nurmi (1999, 2006) contain a collection of such "paradoxes."[20]

These paradoxes tell us that we cannot derive the contribution of an individual to a particular collective action from the individual's voting behavior. Trivially, a vote is not a contribution, but a decision. Resources such as voting power, money, etc. are potential contributions and causality might be traced back to them if collective action results. As a consequence causality follows even from those votes that do not support the collective action. This is reflected in everyday language when one simply states that the Parliament has decided, when in fact decision was made by a majority of less than 100 % of votes. But how can we allocate causality if it cannot be derived from decisions?

[19] This section derives from Holler (2007, 2012) and Holler and Nurmi (2012b).

[20] See also Holler and Nurmi (2012a).

Imagine a five-person committee $N = \{1, 2, 3, 4, 5\}$ that makes a choice between the two alternatives x and y. The voting rule specifies that x is chosen if either (1) 1 votes for x, or (2) at least three of the players 2–5 vote for x. Let us assume that all individuals vote for x. What can be said about causality? Clearly this is a case of over-determination inasmuch as there can be two "winning coalitions" at the same time, and the allocation of causation is not straightforward. The action of agent 1 is an element of only one minimally sufficient coalition, i.e., decisive set, while the actions of each of the other four members are in three decisive sets each. If we take the membership in decisive sets as a proxy for causal efficacy and standardize such that the shares of causation add up to one, then vector

$$h^{\circ} = \left(\frac{1}{13}, \frac{3}{13}, \frac{3}{13}, \frac{3}{13}, \frac{3}{13}\right)$$

represents the degrees of causation.[21] Braham and van Hees (2009: 334), who introduced and discussed the above case, conclude that "this is a questionable allocation of causality." They add that "by focusing on minimally sufficient coalitions, the measure ignores the fact that anything that players 2–5 can do to achieve x, player 1 can do, and in fact more–he can do it alone."

Let us review the above example. Imagine that x stands for polluting a lake. Now the lake is polluted, and all five members of N are under suspicion for having polluted it. Then h° implies that the share of causation for 1 is significantly smaller than the shares of causation of each of the other four members of N. If responsibility and perhaps sanctions follow causation, then the allocation h° seems pathological, at least at the first glance. One might however argue that a smaller member of N could send its garbage to the lake hoping that the lake does not show pollution, while this is not possible in the case of player 1. Given that the costs of cleaning will be assigned to the members of N, player 1's *expected* benefits of sending its garbage to the lake might be much smaller than the expected benefits of the smaller ones.

Perhaps this argument looks somewhat farfetched, but it parallels the "tragedy of the commons" and related "paradoxes" of social interaction. However, Braham and van Hees (2009) propose to apply the weak NESS concept instead of the strong one, i.e., not to refer to decisive sets, but to consider sufficient sets instead and count how often an element i of N is a "necessary element of a sufficient set" (i.e., a NESS).[22] Taking care of an adequate standardization so that the shares add up to 1, we get the following allocation of causation:

---

[21] An alternative measure of "degree of causation" and responsibility is introduced in Chockler and Halpern (2004). It builds on contingency: If a candidate wins an election with 11-0 then a voter who voted for this candidate is less responsible for the victory than if the candidate had won 6-5, but still the voter is responsible under the counterfactual contingency that there could be a 6-5 vote. Similarly, Felsenthal and Machover (2009) allocate responsibility after the decision is made and known.

[22] For a discussion of the NESS test, see Braham (2005, 2008) and Braham and Holler (2009). This literature refers to earlier work by Wright (1985, 1988).

$$b^{\circ} = \left(\frac{11}{23}, \frac{3}{23}, \frac{3}{23}, \frac{3}{23}, \frac{3}{23}\right)$$

The result expressed by $b°$ looks much more convincing than the result proposed by $h°$, does it not? Note that the b-measure and h-measure correspond to the Banzhaf index and the PGI, respectively, and can be calculated accordingly.

If our intuition refers to the capacity of influencing the outcome that differentiates the players, then the numerical results seem to support the weak NESS test and thus the application of the Banzhaf index. However, what happened to alternative y? If y represents "no pollution" then the set of decisive sets consists of all subsets of N that are formed of the actions of agent 1 and the actions of two out of agents 2–5. Thus the actions of 1 are members in six decisive sets while the actions of 2–5 are members of three decisive sets each. The corresponding shares are given by the vector

$$h* = \left(\frac{2}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right)$$

Obviously, $h*$ looks much more convincing than $h°$ and the critical interpretation of Braham and van Hees (2009) no longer applies: agent 1 cannot bring about y on its own, but can cooperate with six different pairs of other agents to achieve this goal.

Note that the actions (votes) bringing about x represent an improper game—two "winning" subsets can co-exist[23] —while the determination of y can be described as a proper game. However, if there are only two alternatives x and y, then "not x" necessarily implies y, irrespective of whether the (social) result is determined by voting or by polluting. The h-values indicate that it seems to matter what issue we analyze and what questions we raise, while the Banzhaf index with respect to y is the same then for x: $b° = b*$.

From the above example we can learn that nonmontonicity might indicate that we asked perhaps the wrong question: Does the responsibility pertain to keeping the lake clean or to polluting it and then perhaps sharing the costs of cleaning it? To conclude, the PGI and thus the strong NESS concept may produce results that are counterintuitive at first glance. However, in some decision situations they seem to tell us more about the power structure and the corresponding causal attribution than the Banzhaf index and the corresponding weak NESS concept do.

In the Republic of San Marino, every six months, the proportionally elected multi-party Council selects two Captains to be the heads of state. These Capitani Reggenti are chosen from opposing parties so that there is a balance of power.

[23] Note that the result x implies the possibility of over-determination. Wright (1985) has identified two types of over-determination: *duplicative* and *pre-emptive* causation. "A case of duplicative causation is one in which two similar and independent causal processes $C1$ and $C2$, each of which is sufficient for the same effect E, may culminate in E at the same time" (Braham and Holler 2009: 149). This applies to x, i.e., the case of pollution.

They serve a six-month term, and a subsequent re-election is not possible. Once their six-month term is over, citizens have three days to file complaints about the Captains' activities. If they are warranted, judicial proceedings against the ex-head(s) of state can be initiated.[24] Should the European Court of Justice evaluate the policies of the Council and the EP? Perhaps impartial commenting could help to make voters more aware of EU decision-making and thus increase political responsibility. However, there have to be more effective ways for the voter to hold his or her representatives accountable than to vote every four years, if responsibility is to work.

Acknowledgments We thank Andreas Nohn and Maurice Salles for their valuable comments.

References

Ahlert, M. & Kliemt, H. (2013). *Necessary and sufficient conditions to make numbers count*. In this volume.

Aleskerov, F., Holler, M. J., & Kamalova R. (2013). *Power distribution in the Weimar Reichstag in 1919–1933*. Annals of Operations Research (forthcoming).

Alonso-Meijide, J. M., & Bowles, C. (2005). Power indices restricted by a priori unions can be easily computed and are useful: a generating function-based application to the IMF. *Annals of Operations Research, 137*, 21–44.

Alonso-Meijide, J. M., Casas-Méndez, B. & Fiestras-Janeiro, M. G. (2013a). *A review of some recent results on power indices*. In this volume.

Alonso-Meijide, J. M., Casas-Méndez, B., Fiestras-Janeiro, M. G., Holler, M. J. & Nohn, A. (2013b). *Coalition configurations and the public good index*. In this volume.

Alonso-Meijide, J. M., & Holler, M. J. (2009). Freedom of choice and weighted monotonicity of power. *Metroeconomica, 60*, 571–583.

Amer, R. & Carreras, F. (2013). *Power, cooperation indices and coalition structures*. In this volume.

Arrow, K. J. (1963). *Social choice and individual values* (2nd ed.). New York: Wiley.

Aumann, R. (1977). Game theory. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *The new Palgrave: a dictionary of economics* (Vol. 2, pp. 460–482). London: MacMillan.

Baldwin, R., & Widgrén, M. (2004). Winners and losers under various dual majority rules for the EU's Council of Ministers. In M. Wiberg (Ed.), *Reasoned choices. Essays in honor of academy professor Hannu Nurmi on the occasion of his 60th birthday* (pp. 42–92). Turku: Digipaino. (also CEPR Discussion Paper No. 4450, June 2004).

Berg, S., & Holler, M. J. (1986). Randomized decision rules in voting games: a model of strict proportional power. *Quality & Quantity, 20*, 419–429.

Bertini, C., Gambarelli, G. & Stach, I. (2013). *Apportionment strategies for the European Parliament*. In this volume.

Braham, M. (2005). Causation and the measurement of power. In G. Gambarelli & M. J. Holler (Eds.), *Power measures III (Homo Oeconomicus 22* (pp. 645–653). Munich: Accedo-Verlag.

[24] The practice of multiple heads of state and the frequent elections of the heads of state are derived directly from the rules that governed the Roman Republic. See Machiavelli's *Discorsi* for an involved analysis of the Roman Republic.

Braham, M. (2008). Social power and social causation: towards a formal synthesis. In M. Braham & F. Steffen (Eds.), *Power, freedom, and voting: Essays in honour of Manfred J. Holler* (pp. 1–21). Berlin and Heidelberg: Springer.

Braham, M. (2013) *Causation and the measurement of power*. In this volume.

Braham, M., & Holler, M. J. (2005a). The impossibility of a preference-based power index. *Journal of Theoretical Politics, 17*, 137–157.

Braham, M., & Holler, M. J. (2005b). Power and preferences again: a reply to Napel and Widgrén. *Journal of Theoretical Politics, 17*, 389–395.

Braham, M., & Holler, M. J. (2009). Distributing causal responsibility in collectivities. In R. Gekker & T. Boylan (Eds.), *Economics, rational choice and normative philosophy* (pp. 145–163). London and New York: Routledge.

Braham, M., & van Hees, M. (2009). Degrees of causation. *Erkenntnis, 71*, 323–344.

Carreras, F. & Owen, G. (2013). *Pure bargaining problems and the Shapley value*. In this volume.

Carter, I. (2013). *Social power and negative freedom*. In this volume.

Chockler, H. & Halpern, J.Y. (2004). Responsibility and blame: A structural-model approach. *Journal of Artificial Intelligence Research*, *22*, 93–115.

Coleman, J. (1971). Control of collectivities and the Power of a collectivity to act. In B. Lieberman (Ed.), *Social choice*. New York: Gordon and Breach.

Deegan, J, Jr, & Packel, E. W. (1979). A new index of power for simple n-person games. *International Journal of Game Theory, 7*, 113–123.

De Parle, J. (2007). The American prison nightmare. *The New York Review of Books 56 (April 12)* (pp. 33–36).

Eckert, D. & Klamler, C. (2013). *Geometric approach to paradoxes of majority voting: from Anscombe's paradox to discursive dilemma with Saari and Nurmi*. In this volume.

Felsenthal, D., & Machover, M. (1998). *The measurement of voting power. Theory and practice, problems and paradoxes*. Cheltenham: Edward Elgar.

Felsenthal, D., & Machover, M. (2001). The treaty of nice and qualified majority voting. *Social Choice and Welfare, 18*, 431–464.

Felsenthal, D., & Machover, M. (2009). A note on measuring voters' responsibility. *Homo Oeconomicus, 26*, 259–271.

Felsenthal, D. S. & Machover, M. (2013a). *Models and reality: the curious case of the absent abstention*. In this volume.

Felsenthal, D. S. & Machover, M. (2013b). *The QM rule in the Nice and Lisbon treaties: future projections*. In this volume.

Felsenthal, D., Machover, M., & Zwicker, W. (1998). The bicameral postulates and indices of a Priori voting power. *Theory and Decision, 44*, 83–116.

Freixas, J. & Pons, M. (2013). *Circumstantial power: some hints for finding optimal persuadable or bribable voters*. In this volume.

Gambarelli, G., & Owen, G. (1994). Indirect control of corporations. *International Journal of Game Theory, 23*(4), 287–302.

Gambarelli, G., & Owen, G. (2002). Power in political and in business structures. In M. J. Holler, H. Kliemt, D. Schnidtchen, & M. E. Streit (Eds.), *Power and fairness* (pp. 57–68). Tübingen: Mohr Siebeck. (Jahrbuch für Neue Politische Ökonomie 20).

Gambarelli, G. & Palestini, A. (2013). *Minimax multi-district apportionments*. In this volume.

Garrett, G., & Tsebelis, G. (1996). An institutional critique of intergovernmentalism. *International Organization, 50*, 269–299.

Garrett, G., & Tsebelis, G. (1999). Why resist the temptation to apply power indices to the EU. *Journal of Theoretical Politics, 11*, 291–308.

Hart, S., & Mas-Colell, A. (1996). Bargaining and value. *Econometrica, 64*, 357–380.

Holler, M. J. (Ed.). (1982a). *Power, voting, and voting power*. Würzburg and Wien: Physica-Verlag.

Holler, M. J. (1982b). An introduction to the analysis of power, voting, and voting power. In M. J. Holler (Ed.), *Power, voting, and voting power* (pp. 15–30). Würzburg and Wien: Physica-Verlag.

Holler, M. J. (1982c). Forming coalitions and measuring voting power. *Political Studies, 30*, 262–271.

Holler, M. J. (1984). The public good index. In M. J. Holler (Ed.), *Coalitions and collective action*. Würzburg and Vienna: Physica-Verlag.

Holler, M. J. (1985). Strict proportional power in voting bodies. *Theory and Decision, 19*, 249–258.

Holler, M. J. (Ed.). (1987a). *The logic of multiparty systems*. Dordrecht: Kluwer Academic Publishers.

Holler, M. J. (1987b). Paradox proof decision rules in weighted voting. In M. J. Holler (Ed.), *The logic of multiparty systems*. Dordrecht: Kluwer Academic Publishers.

Holler, M. J. (1998). Two stories, one power index. *Journal of Theoretical Politics, 10*, 179–190.

Holler, M. J. (2007). Freedom of choice, power, and the responsibility of decision makers. In J.-M. Josselin & A. Marciano (Eds.), *Democracy, freedom and coercion: a law and economics approach* (pp. 22–45). Cheltenham: Edward Elgar.

Holler, M. J. (2012). EU decision-making and the allocation of responsibility. In T. Eger & H.-B. Schäfer (Eds.), *Research handbook on the economics of European Union Law* (pp. 55–81). Cheltenham: Edward Elgar.

Holler, M. J., & Napel, S. (2004a). Local monotonicity of power: axiom or just a property. *Quality & Quantity, 38*, 637–647.

Holler, M. J., & Napel, S. (2004b). Monotonicity of power and power measures. *Theory and Decision, 56*, 93–111.

Holler, M. J., & Nurmi, H. (2010). Measurement of power, probabilities, and alternative models of man. *Quality & Quantity, 44*, 833–847.

Holler, M. J., & Nurmi, H. (2012a). Aspects of power overlooked by power indices. In R. Fara, D. Leech, & M. Salles (Eds.), *Voting power and procedures. Essays in honour of Dan Felsenthal and Moshé Machover*. Berlin et al: Springer (forthcoming).

Holler, M. J., & Nurmi, H. (2012b). Pathology or revelation?—the public good index. In R. Fara, D. Leech, & M. Salles (Eds.), *Voting power and procedures. Essays in honour of Dan Felsenthal and Moshé Machover*. Berlin et al: Springer (forthcoming).

Holler, M. J., & Wegner, W. (2011). Power and responsibility in environmental policy making. *AUCO Czech Economic Review, 5*, 267–288.

Holler, M. J., & Widgrén, M. (1999a). Why power indices for assessing European Union decision-making? *Journal of Theoretical Politics, 11*, 321–330.

Holler, M. J., & Widgren, M. (1999b). The value of a coalition is power. *Homo Oeconomicus, 15*, 485–501. (reprinted in *Notizie di Politeia* 59, 2000, 14–28).

Holler, M. J., Ono, R., & Steffen, F. (2001). Constrained monotonicity and the measurement of power. *Theory and Decision, 50*, 385–397.

Johnston, R. (1978). On the measurement of power: some reactions to laver. *Environment and Planning, 10A*, 907–914.

Johnston, R. (2013). *Power indices and the design of electoral/constitutional systems*. In this volume.

Kauppi, H., & Widgrén, M. (2004). What determines EU decision making? Needs, power or both? *Economic Policy, 19*, 221–266.

Kirsch, W. (2013). *On Penrose's square-root law and beyond*. In this volume.

König, T. & Bräuninger, T. (2013). Decisiveness and inclusiveness: two aspects of the intergovernmental choice of the European voting rules. In this volume.

Kauppi, H., & Widgrén, M. (2007). Voting rules and budget allocation in an enlarged EU. *European Journal of Political Economy, 23*, 693–706.

Laruelle, A., & Valenciano, F. (2001). Shapley-Shubik and Banzhaf indices revisited. *Mathematics of Operations Research, 26*, 89–104.

Laruelle, A., & Valenciano, F. (2008). Noncooperative foundations of bargaining power in committees and the Shapley-Shubik index. *Games and Economic Behavior, 63*, 341–353.

Laruelle, A. & Valenciano, F. (2013). *Voting and power*. In this volume.

Laruelle, A., & Widgrén, M. (1998). Is the allocation of voting power among EU states fair? *Public Choice, 94*, 317–338.

Leech, D. (2013). *Shareholder voting power and ownership control of companies*. In this volume.

Leech, D. & Leech, R. (2013). *A new analysis of a priori voting power in the IMF: recent quota reforms give little cause for celebration*. In this volume.

Montero, M. (2006). Noncooperative foundations of the nucleolus in majority games. *Games and Economic Behavior, 54*, 380–397.

Montero, M. (2013). *On the nucleolus as a power index*. In this volume.

Napel, S., & Widgrén, M. (2001). Inferior players in simple games. *International Journal of Game Theory, 30*, 209–220.

Napel, S., & Widgrén, M. (2005). The possibility of a preference-based power index. *Journal of Theoretical Politics, 17*, 377–387.

Napel, S., & Widgrén, M. (2006). The inter-institutional distribution of power in EU codecision. *Social Choice and Welfare, 27*, 129–154.

Napel, S., Widgrén, M. & Mayer, A. (2013). *Strategic a priori power in the European Union's codecision procedure recalculated for EU28*. In this volume.

Nohn, A. (2013). *Veto players and non-cooperative foundations of power in legislative bargaining*. In this volume.

Nurmi, H. (1999). *Voting paradoxes and how to deal with them*. Berlin, Heidelberg and New York: Springer.

Nurmi, H. (2006). *Models of political economy*. London and New York: Routledge.

Ono-Yoshida, R. (2013). *Limit property of a multi-choice value and the fuzzy value*. In this volume.

Owen, G. (1977). Values of games with a priori unions. In R. Henn & O. Moeschlin (Eds.), *Mathematical economics and game theory*. Berlin and New York: Springer. (Lecture notes in economics and mathematical systems 141).

Owen, G. (1982). Modification of the Banzhaf-Coleman index for games with a priori unions. In M. J. Holler (Ed.), *Power, voting, and voting power*. Würzburg and Wien: Physica-Verlag.

Parsons, T. (1928). 'Capitalism' in recent German literature: Sombart and Weber. *Journal of Political Economy, 36*, 641–661.

Parsons, T. (1929). 'Capitalism' in recent German literature: Sombart and Weber. *Journal of Political Economy, 37*, 31–51.

Penrose, L. S. (1946). The elementary statistics of majority voting. *Journal of the Royal Statistical Society, 109*, 53–57.

Pérez-Castrillo, D., & Wettstein, D. (2001). Bidding for the surplus: a non-cooperative approach to the Shapley value. *Journal of Economic Theory, 100*, 274–294.

Rapoport, A. (1982). Foreword. In M. J. Holler (Ed.), *Power, voting, and voting power* (pp. 5–7). Würzburg and Wien: Physica-Verlag.

Schofield, N. (2009). *The political economy of democracy and tyranny*. Munich: Oldenbourg-Verlag.

Schofield, N. (2013). *Coalitions and catastrophic climate change*. In this volume.

Shapley, L. S., & Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review, 48*, 787–792.

Słomczyński, W., & Życzkowski K. (2007b). Jagiellonian compromise: an alternative voting system for the council of the European Union. http://chaos.if.uj.edu.pl/~karol/pdf/JagCom07.pdf.

Słomczyński, W., & Życzkowski, K. (2006). Penrose voting system and optimal quota. *Acta Physica Polonia B, 37*, 3133–3143.

Słomczyński, W., & Życzkowski, K. (2007). From a toy model to the double square root voting system. *Homo Oeconomicus, 24*, 381–399.

Steunenberg, B., Schmidtchen, D., & Koboldt, C. (1999). Strategic power in the European Union: evaluating the distribution of power in policy games. *Journal of Theoretical Politics, 11*, 339–366.

Sutter, M. (2000). Fair allocation and re-weighting of votes and voting power in the EU before and after the next enlargement. *Journal of Theoretical Politics, 12*, 433–449.

Tsebelis, G., & Garrett, G. (1996). Agenda setting power, power indices, and decision making in the European Union. *International Review of Law and Economics, 16*, 345–361.

Tsebelis, G., & Garrett, G. (1997). Why power indices cannot explain decision making in the European Union. In D. Schmidtchen & R. Cooter (Eds.), *Constitutional law and economics of the European Union*. Cheltenham: Edward Elgar.

Turnovec, F. (2004). Power indices: swings or pivots? In M. Wiberg (Ed.), *Reasoned choices. Essays in honour of Academy Professor Hannu Nurmi on the Occasion of his 60th birthday*. Turku: Digipaino.

Turnovec, F. (2013). *Fair voting rules in committees*. In this volume.

Turnovec, F., Mercik, J., & Mazurkiewicz, M. (2008). Power indices methodology: deciveness, pivots, and swings. In M. Braham & F. Steffen (Eds.), *Power, freedom and voting. Essays in honour of Manfred J Holler* (pp. 23–37). Berlin and Heidelberg: Springer.

Weber, M. (1947[1922]). The theory of social and economic organization. Trans. by A.M. Henderson and T. Parsons. New York: Free Press.

Weber, M. (1948[1924]). Class, status and party. In H. H. Gerth & C. W. Mills (Eds.), *Essays from Max Weber*. London: Routledge and Kegan Paul.

Weber, M. (2005[1922]). *Wirtschaft und Gesellschaft. Grundriß der verstehenden Soziologie*. Zweitausendeins: Frankfurt am Main.

Widgrén, M. & Napel, S. (2013) *The power of spatially inferior player*. In this volume.

Wright, R. (1985). Causation in tort law. *California Law Review, 73*, 1735–1828.

Wright, R. (1988). Causation, responsibility, risk, probability, naked statistics, and proof: pruning the bramble bush by clarifying the concepts. *Iowa Law Review, 73*, 1001–1077.

Yan, H. (2002). Noncooperative selection of the core. *International Journal of Game Theory, 31*, 527–540.

Życzkowski, K., & Słomczyński, W. (2013). *Square root voting system, optimal threshold and π*. In this volume.

Saari, D. (2013). Explaining all possible paired comparisons. In this volume.

Vartiainen, H. (2013). Distortion-free logrolling mechanism. In this volume.

Part I
Power

Social Power and Negative Freedom

Ian Carter

1 Introduction

When agent A exercises power over agent B, what is the effect on B's freedom? Is B less free as a result? Does A remove any specific freedoms of B? Most of us feel intuitively that there are many kinds of social power, and that while some of these may affect B's freedom to a great extent, others may affect it less, and others still may leave B's freedom completely intact. It would seem to be important for philosophers and social scientists to provide an explicit and coherent explication of this intuitive relation between the social power of A and the unfreedom of B. Nevertheless, surprisingly little attention has so far been devoted to its analysis.

One reason for the relative lack of interest in the freedom-power relation may lie in the different theoretical outlooks dominant within the disciplinary areas within which these two concepts tend to be examined and applied. The concept of freedom has been analyzed above all by political philosophers interested in its role within normative theories and thus in its relation to concepts like equality, justice, toleration, rights and the rule of law. Power, on the other hand, is a fundamental concept in the social sciences, where little attention has been devoted to the concept of freedom. Political scientists often express the view, shared by a number

This article was previously published in *Homo Oeconomicus* 24 (2007), pp. 187–229. I am grateful to Francesco Battegazzorre, Giorgio Fedel and Laura Valentini for comments on a previous, Italian version of this article, and to the participants in the Colloquium on Freedom held at the University of Bayreuth in February 2005.

I. Carter (✉)
Department of Political and Social Sciences, University of Pavia, Pavia, Italy
e-mail: icarter@unipv.it

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_2,

of influential political philosophers, that freedom is an irreducibly evaluative term, ill-suited to empirical research and theorizing.[1]

This is something of a shame, for there is nevertheless a strong current of thought within contemporary political philosophy according to which we have a theoretical interest in conceiving of freedom in purely empirical terms. Various reasons have been advanced in defence of this stance. One such reason is that our understanding of normative disagreements about freedom is best furthered by our first establishing *agreement* over who is free and who unfree (or who is free to what extent) and then investigating how people disagree in their *evaluations* of these agreed facts about freedom (Oppenheim 1961, 1981; Steiner 1994; Kramer 2003). Another reason, which is more internal to liberal political theory, concerns the role of freedom as a fundamental value: if freedom is a fundamental value, it provides a reason for our wishing to promote certain other, less fundamental values, in which case it will not do to define freedom *in terms of* those other values. Instead, it is argued, freedom should be defined in terms that are independent of those values (Cohen 1991). Yet another reason (again a liberal one) is that freedom has a special kind of value which may be called 'non-specific'. If freedom is non-specifically valuable, then its value is not wholly constituted by the value of being free to do one or another specific thing or set of things, for freedom also has value *as such*. Elsewhere I have contended that a purely empirical measure of freedom is needed to capture the sense we have of freedom being non-specifically valuable (i.e., valuable as such) (Carter 1999; cf. van Hees 2000).

As a political philosopher, my own reason for taking an interest in the freedom-power relation is that liberals often wish to condemn certain forms of power, or certain distributions of those forms of power, because of their effects on freedom. They also aim, on this basis, to construct normative political theories—including models of political institutions—that limit power or that distribute power in a certain way (or that do both of these things) in the name of freedom. This is especially true of contemporary republican political theory.

While my own reason for investigating the freedom-power relation is a normative one, however, my analysis ought not to be of interest only to normative political theorists. For the kind of relation it will posit between these two phenomena is an empirical relation. If freedom and power are both understood as empirical, explanatory phenomena, a plausible theory about how they are related might well be of interest to social scientists—just as, say, a plausible theory about the relation between electoral systems and political stability ought to be of interest to them.

I shall take as my starting point a particular 'negative' conception of freedom that I have already defended elsewhere (Carter 1999), and the formal classification of social power originally set out by Stoppino (2007—see also Table 1). There are at least two good reasons for taking Stoppino's classification as a fixed point of

[1] Economists have recently begun to show interest in the concept of freedom—especially in the area of social choice theory—but have yet to turn their attention systematically to its relation to power. An exception is Braham (2006), but this is not concerned with the different forms of power, in the sense of 'form' I shall assume in this article.

Table 1 Stoppino’s formal classification of power

<table>
<tr><td colspan="2">target of intervention
open/hidden
intentional/interested</td><td>social environment</td><td>unconscious psychological processes</td><td>factual knowledge and value beliefs</td><td>available alternatives</td></tr>
<tr><td>hidden power</td><td rowspan="2">intentional and interested</td><td>situational manipulation</td><td>psychological manipulation</td><td>informational manipulation</td><td></td></tr>
<tr><td rowspan="2">open power</td><td>conditioning</td><td></td><td>persuasion</td><td>coercion
remuneration</td></tr>
<tr><td>merely interested</td><td>merely interested conditioning</td><td></td><td>imitation</td><td>anticipated reactions</td></tr>
</table>

reference for this investigation. First, it is a fine-grained classification, and will therefore allow us to distinguish between the effects on freedom of a suitably large number of forms of power. Secondly, his classification will lend clarity to the freedom-power relation through the implied additional distinction between forms of power and the substantive means by which those forms can be exercised. These substantive means consist in instruments of violence, economic resources, and symbolic resources. Their different effects on freedom will be taken into account in my analysis, but it is important to maintain the distinction between the effects on freedom of the uses of these different substantive means and the effects on freedom of different forms of power like, for example, coercion, remuneration and manipulation. Other typologies of power have involved slippage between the formal and substantive categories, resting on distinctions such as that between coercive power and economic power. For Stoppino, plausibly enough, coercive power and economic power are not mutually exclusive: economic resources are just one of the means by which coercive power, remunerative power, conditioning, and so on, may be exercised. (This said, my analysis of the freedom-power relation will also imply some minor criticisms of Stoppino’s classification, regarding both the definitions of the forms of power and the collocation of some of his examples).

Although I assume a particular negative conception of freedom here, my investigation is not intended primarily as a polemic against those who assume rival conceptions; its central aim is simply to clarify the relation between two concepts. Nevertheless, I hope that the intuitive plausibility of the results of the analysis will serve to strengthen the case for the conception of freedom it assumes.

2 Violence, Preclusion and Freedom

A fundamental distinction made by Stoppino is that between power and violence. Although everyday discourse assumes the use of violence to be a form of power, for Stoppino this is not the case. Power, understood as a social relation, consists in the modification by A of B's conduct (or the possibility for A of bringing about that modification) in A's interests, where the expression 'B's conduct' refers to an action or omission (or set of actions or omissions) that is voluntary, at least to a minimal degree. If A exercises power over B, A modifies B's behaviour by means of an intervention on B's *will*, such that, while in the absence of A's intervention B would have done *x*, in the presence of that intervention B decides not to do *x*. Violence, on the other hand, is a physical intervention on the part of one agent directly on the body or the immediate physical environment of another. If A behaves violently towards B, A modifies B's behaviour directly rather than by means of B's will, preventing B's doing *x* by physically removing that option. When A brings about the same behaviour through power over B, on the other hand, A does not remove B's option of doing *x*, but instead brings it about that B decides not to do *x*. Thus, if A holds a gun to B's head and tells B to leave the room, as a result of which B leaves the room under his own steam, then A exercises power over B. But if A physically pushes B out of the room, A is simply engaging in violent behaviour. The agent who exercises power does so through 'persuasion, the threat of punishment, the promise of a reward, the appeal to authority, setting an example, the rule of anticipated reactions, and so on' (Stoppino 2001a, p. 73), whereas the violent agent is one 'who attacks, wounds or kills; who, notwithstanding any resistance, immobilizes or manipulates the body of another; who materially prevents another from performing a certain action' (Stoppino 2001a, p. 70).

In what follows, I shall contrast power relations not only with violent relations but also with the wider category of *preclusive* relations, of which violent relations are a sub-category. To see the difference between violence and mere preclusion, consider the following example in which the Australian government (A), partly determines the behaviour of a permanent resident of Milan (B). Imagine that the Australian government fences off the entire Australian outback, thus precluding entry by the Milan resident. Regardless of whether or not the Milan resident was in fact planning a visit to the Australian outback, the Australian government's intervention physically determines the fact that the Milan resident does not enter the outback. The physical determination of this fact about the Milan resident's behaviour does not constitute an exercise of power over the Milan resident, as the fact of the Milan resident not entering the outback is not (after the erection of the fence) a product of the Milan resident's *will*. But neither is the intervention plausibly described as one of violence, for it is not an intervention on the body or the *immediate* physical environment of the Milan resident. The intervention is simply one of preclusion. Violence is only one kind of preclusion, although the most invasive kind: being an intervention on the agent's body or *immediate* physical environment, it tends to preclude a great deal. The importance of contrasting power not only with violence but also with preclusion more generally will become clear later on, when we come to examine the relation between freedom and manipulation.

The conception of freedom I shall assume here is often called ‘pure negative’ freedom. (On the distinction between ‘negative’ and ‘positive’ freedom, see Berlin 2002; Carter 2003; Carter et al. 2007). According to this conception, unfreedom is a social relation consisting in the presence of humanly imposed impediments rendering actions *impossible*. Social impediments to action that do not render actions impossible—for example, physical obstacles that can be overcome at great cost or pain—do not render the agent unfree to perform those actions. Instead, what they do is render those actions more costly or painful. Thus, if my neighbour were to erect a three-metre wall around his garden—the kind of wall that I am simply unable to scale, even with the greatest effort—I would be unfree to enter my neighbour’s garden. But if the wall were only two metres high, and I were able to scale it with a huge amount of effort, then I would be free to enter the garden.[2] Similarly, when an agent is deterred from doing *x* by the prospect of costs that would be incurred subsequent to her doing *x*, that agent is nevertheless free to do *x*. Thus, I would be free to enter the garden if (a) my neighbour offered to open a door in the wall but only on condition that I sign over to him my entire salary for the next three years, or (b) my neighbour opened the door but issued a credible threat to kill me should I ever pass though it.

There are various reasons for assuming this conception of freedom (Taylor 1982; Gorr 1989; Steiner 1994; Carter 1999; Kramer 2003), despite the initial doubts that are often provoked by examples like those I have just cited. Here, I shall mention one such reason that is particularly salient in the context of the freedom-power relation: the threat to punish agent B for doing *x* does not remove B’s freedom to do *x* for the same reason that the offer to reward B for doing not-*x* does not remove B’s freedom to do *x*. As Hillel Steiner has argued (and as Stoppino implicitly agrees), the *modus operandi* of an offer is not different from that of a threat: both interventions invert the preference order of the agent with respect to the alternatives of doing *x* and not doing *x* (Steiner 1994, Chap. 2). The fact that a threat works by reducing the desirability of *x*, whereas an offer works by increasing the desirability of not-*x*, is not a relevant difference when it comes to estimating the degree of *effectiveness* of an intervention in bringing it about that the agent does not-*x*. That degree of effectiveness depends only on the size of *the difference in desirability* (for B) between *x* and not-*x* that the intervention is able to bring about. For example, the offer to reward B with $10,000 for forbearing from parking her car a certain space will normally be a much more powerful intervention than the threat to fine B $10 for parking there. It will be more likely to succeed in inverting the preference order of B with respect to parking and not parking, because it raises the value of not-x much more than the threat lowers the value of *x*.

[2] For the sake of simplicity, I here assume that freedom is the absence of unfreedom, so that ‘not unfree’ entails ‘free’ (and ‘not free’ entails ‘unfree’). This bivalence assumption is not unproblematic, but I shall not discuss the issue here. For a critique, see Kramer 2003, pp. 41–60.

Were we to say that a threat against doing *x* removes the agent's freedom to do *x*, then, we should have to say the same of an equally powerful offer (and, *a fortiori*, of a more powerful offer) to reward the agent for doing not-*x*. Yet it is highly counterintuitive, from the liberal point of view that favours a so-called 'negative' conception of freedom, to say that offers restrict freedom, for in order to say this we should have to make certain assumptions about freedom that are more characteristic of so-called 'positive' conceptions—for example, that freedom consists, at least in part, in self-direction, or in autonomy of the will.[3] It is for this reason that the conception that rules out threats and offers as sources of unfreedom is called the 'pure' negative conception.

Now, from the pure negative conception of freedom it follows that there is no connection between an agent's negative freedom and her *will* (even though, of course, her freedom depends on the wills of *other* agents to act in certain ways rather than others). Only positive or 'impure' negative conceptions of freedom allow the state of B's will to affect the question of whether or how far B is free. On the pure negative conception, I am unfree to do *x* if and only if someone else has rendered *x* impossible for me, regardless of whether or not I want to do *x*. This fact, however, might be thought to give rise to a problem: the lack of connection between B's freedom and B's will, in conjunction with Stoppino's insistence on the nature of power as mediated by B's will, would seem to suggest that A's power over B *never* affects the freedom of B. We have seen that power exercised by A over B necessarily presupposes a minimum of voluntariness on the part of B. It presupposes B's possibility of doing otherwise. Rendering an action impossible, on the other hand, removes that minimum of voluntariness, and is therefore at most an instance of *preclusion*. Thus, all instances of A restricting B's freedom would appear to fail to qualify as instances of A exercising power over B. Is it not sheer common sense, however, to say that the freedom of one agent depends on an absence of at least certain kinds of power on the part of other agents? Are not the power of A and the unfreedom of B, at least to some extent, two sides of the same coin?

The analysis presented in this article will show the above dilemma to be illusory: we need not choose between the pure negative conception of freedom and the tendency to associate the power of A with the unfreedom of B. Indeed, one of my central aims is to show how, even on the pure negative conception of freedom, B's freedom is restricted by a number of different forms of power on the part of A. Two distinctions within the concept of freedom will be central to the pursuit of this aim. The first is the distinction between 'the freedom to act' and 'acting freely', and will be applied in the next section. The second is the distinction between specific freedoms and overall freedom, and will be applied in the subsequent section.

Before starting, two preliminary points should be made. First, I shall take for granted that there are cases of 'power without unfreedom'. No one who endorses a

---

[3] In Carter 2008, I apply this observation to an analysis of Philip Pettit's notion of freedom as "discursive control" (Pettit 2001), arguing that freedom as discursive control is limited by offers, no less than by threats.

negative conception of freedom (whether pure or 'impure') claims that freedom depends on the absence of power *tout court*, unless the notion of power is understood in a much narrower sense than that of Stoppino. For example, while rational persuasion is a form of power on Stoppino's analysis, no supporter of a negative conception of freedom would say that when A rationally persuades B to do *x*, A somehow renders B socially unfree.

Secondly, I shall similarly take for granted that there are cases of 'unfreedom without power'. Thus, while on the first assumption I have just mentioned A's power is not a sufficient condition for B's unfreedom, on this second assumption it is not a necessary condition either. The question I am asking myself in this article is not whether, or how far, restrictions of freedom are the result of power relations, but whether, or how far, power relations result in restrictions of freedom. In other words, I am not asking whether unfreedom implies power, but whether power implies unfreedom. We have already seen that there are cases of unfreedom that are caused not by power relations but by intentional or interested preclusion (including violence). Virtually no one would deny that there are *some* such cases. In addition to these cases, we should also count as ones of 'unfreedom without power' those in which A's behaviour precludes B's doing *x* but in a way that is neither violent nor intentional nor in A's interests. As Stoppino would put it, in the latter cases the relation between A and B is neither of violence nor of power, because each of these two kinds of relation necessarily involves A's 'interested' modification of B's behaviour. Many theorists of negative freedom—among whom the supporters of the pure negative conception—would nevertheless say that in all such cases of preclusion, A restricts B's freedom. Pure negative unfreedom is normally conceived as the result of obstruction by other agents, regardless of whether that obstruction is intentional or unintentional, interested or disinterested. Power and unfreedom are therefore asymmetrical in this respect: while A's power over B depends on a furthering of A's interests, A's restriction of B's freedom does not.

3 Power and Acting Freely

The freedom *to act*, understood in the negative sense outlined above, consists in the absence of constraints on an agent's possible actions. One's freedom to act is, to use Isaiah Berlin's metaphor, a matter of how many doors are open to one (Berlin 2002, pp. 32, 35), and one's particular conception of the freedom to act will depend, among other things, on how one defines the closing of a door. The freedom *of an action*, on the other hand, is to be found in the *performance* of that action. The freedom of one's actions—i.e., whether or how far one *acts freely* when one does act—is therefore a question not so much of how many doors are open as of how and why one goes through one door rather than another. Appropriating (and slightly modifying) a distinction introduced by Charles Taylor, we can say that whereas the concept of freedom to act is an 'opportunity concept', the concept of acting freely is an 'exercise concept', given that the latter concerns the way in which a certain possibility is realized or exercised (Taylor 1979). Oppenheim has clarified this distinction by noting that while in the first case (the freedom to act) freedom is a *property of an agent*, in the second case (free action or acting freely) freedom is a *property of an action* (Oppenheim 1981, Sect. 5.2).[4] Oppenheim himself stipulates that an action is performed unfreely it if is performed out of fear of a sanction. Others have suggested broader definitions of acting unfreely. According to Serena Olsaretti, for example, an action is performed unfreely if the reason for its performance is that the agent has no acceptable alternative (Olsaretti 2004, Chap.6).

Like the concept of freedom, the concept of social power can be interpreted either as an opportunity concept or as an exercise concept.[5] On the one hand, one can *have* power, in the sense of having the option of modifying the conduct of another in one's own interests. Here, power is an opportunity concept, which Stoppino calls 'potential power'. On the other hand, one can *exercise* power, in the sense of bringing about that modification in the conduct of another. Here, power is an exercise concept, which Stoppino calls 'actual power'. In this section and the next, I shall assume that the kind of power of A we are concerned with, in discussing the implications for B's freedom, is A's actual power—power A *exercises* over B. It is also true, however, that the potential power of A can limit B's freedom even without A exercising that power, as long as there is some probability of A exercising it (Carter 2008). I shall come to the role of probabilities in the next section.

The distinction between the freedom to act and acting freely is present in Stoppino's analysis of power. We have seen that for Stoppino, when A exercises power over B, B's behaviour is always characterized by a minimum degree of voluntariness, such that B could have done otherwise. However, Stoppino does not believe that B's action is for this reason 'free' (Stoppino 2001a, p. 6, 73). If a bandit says to me 'Your money or your life', and I hand over my money for fear of being killed, I do so voluntarily in the sense that I could have refused to hand over the money and borne the consequences of the bandit's subsequent violent intervention. Nevertheless, we tend to think that my choice to hand over the money is nevertheless not a 'free' choice, because the reason behind the choice consists in fear of a severe sanction. These two views are not mutually exclusive. While my choice is not 'freely taken', it remains the case that I could have done otherwise, had I so desired: my behaviour is voluntary in the minimal sense of having my own will as its proximate cause.

This voluntariness, in Stoppino's sense of the term 'voluntariness', stands for what I would call a freedom of the agent *to act*: the agent who is subject to coercive power is free not only to comply with the threat but also to refuse compliance. As long as we bear in mind the distinction between the freedom to act and acting freely, then, we can reasonably attribute to Stoppino not only the view that those who are coerced into doing *x* remain *free* not to do *x*, but also the view

---

[4] It should be added, however, that one may also go on to predicate freedom of agents (in the exercise sense) on the basis of the fact that they perform their actions freely.

[5] Unlike the exercise concept of power, the exercise concept of acting freely is not necessarily a concept of *social* freedom. For example, of the two definitions just mentioned, Oppenheim's concept of acting freely is a social concept, but Olsaretti's is not.

that A's exercising power over B often brings about a *certain* kind of *unfreedom*, namely the unfreedom with which B actually does x.

There is indeed a difference between the freedom to act and acting freely that makes the connection with social power much more immediate and obvious in the case of the latter concept than in that of the former. This difference lies in the fact that my acting freely or unfreely in doing x depends on my *motivation* for doing x—for Oppenheim, it depends on whether I act out of fear of a sanction; for Olsaretti, it depends on whether my reason for doing x is that I lack any acceptable alternative. On the other hand, it is plausible to claim (and we have seen that defenders of the pure negative conception do indeed claim) that the question of whether I am free *to* do x does *not* depend in any way on my motivational state. In this sense, the presence of A's coercive power over B points much more obviously to the fact of B *acting unfreely* than to any unfreedom of B *to act*. The fact of my being subject to the power of another clearly *depends on me*, in the sense of depending on *how my own will reacts* to that of another, whereas the same is not true of my being subject to a social unfreedom to act.

It is easy enough to confuse acting freely with the freedom to act. Joseph Goebbels confused them when he claimed, ironically, that 'anyone is free to write what he likes as long as he is not afraid of the concentration camp' (cited in Gabor and Gabor1979, p. 346). This claim is not literally false, but it is confused, or at least confusing, because it can reasonably be taken to imply the further claim that anyone who *is* afraid of the concentration camp is *not* free to write what he likes, and the latter claim *is* false. In Nazi Germany, the freedom to write what one likes (up until the moment of arrest) was possessed both by those who were not afraid of the concentration camp and by those who were afraid of the concentration camp. On the other hand, there is a difference between these two classes of people in terms of how freely they chose *not* to express their views in writing (where they did so choose). Assuming Oppenheim's definition of free action, we should say that those who were afraid of the concentration camp chose unfreely to avoid expressing their views in writing, whereas those who were not afraid of the concentration camp suffered no restriction on the freedom with which they chose not to express their views in writing (they would have chosen not to do so even in the absence of Goebbels' threat).

One reason for the ease of slippage between the concepts of freedom to act and acting freely lies in an ambiguity in the term *voluntariness*. This ambiguity is mirrored by the different technical meanings attributed to the term in the literature, some authors taking it to signify the presence of a freedom to act, others the fact of acting freely. For Stoppino, as well as for Oppenheim (1981, Sect. 5.2), those who comply with a coercive threat act 'voluntarily', in the sense of having been *free to act differently*. Here, the voluntariness of an action signifies no more than that its proximate cause is the will of the agent.[6] For some theorists of freedom, however,

[6] The English terms 'will' and 'voluntariness' have different etymological roots. The connection between them is much clearer in Latin languages (their respective equivalents in Stoppino's native tongue are *volontà* and *volontarietà*).

those who comply with a coercive threat act in a non-voluntary way, because voluntary action is conceived by them as identical to what I have so far referred to as the fact of *acting freely* (this is the case, for example, in Olsaretti 2004, and in van Hees 2003). The difference between these two sets of authors is clearly terminological rather than substantive. The important point to bear in mind, for present purposes, is that the pure negative conception of freedom is a conception of freedom to act, not of acting freely, and that it does not conflict at all with the claim that power creates unfreedom in the sense of leading people to act unfreely.

4 Power and the Freedom to Act

Admitting that power restricts the freedom with which people act will not, however, be sufficient to allay the worries of those who initially see the pure negative conception of freedom as unable to capture the freedom-restricting effects of power. For among the sources of such worries one must certainly count the intuition that when A exercises power over B, A limits B's *freedom to act*. Is it possible for the supporter of the pure negative conception to accommodate this intuition too? I believe that it is. In order to show how, we shall need now to make a distinction within the concept of the freedom to act: that between a specific freedom and overall freedom (Carter 1999, Chap. 1).

A specific freedom is the freedom of an agent to perform a specific action—for example, my freedom to leave this room in ten minutes' time (a freedom that I shall lose if, during the next ten minutes, someone locks the door). I shall assume here that by 'specific freedom' we mean the freedom to perform a *spatio-temporally* specific action—not a specific *type* of action (such as walking or talking), but a *concrete particular*, unrepeatable both in time and in space (like the freedom to move out of *this* room in exactly *ten minutes' time*). Overall freedom, on the other hand, is a quantitative attribute of an agent. It is the freedom the agent possesses *in a certain degree*. Overall freedom is still the freedom *to act*, but it is not the freedom to perform some *specific* action. Instead, it consists in an aggregation of all the agent's freedoms and unfreedoms, so providing us with an overall quantitative judgement about the extent to which the agent is free to act (be this in absolute terms or only relative to the extents of freedom of other agents). The possibility of forming coherent quantitative judgements about overall freedom is presupposed whenever one agent, group or society is described as 'more free' than another, whenever it is claimed that citizens have a right to 'equal freedom', and whenever theorists or politicians prescribe that freedom in society, or freedom for certain groups, be 'increased', 'augmented', 'maximized', or maintained above 'a certain minimum'.

4.1 The Non-Specific Values of Freedom and Power

Before turning to the effect of power on overall freedom, it is worth noting a parallel between the kinds of value attributed to power and freedom that motivate an interest not only in the concept of overall freedom but also in that of overall power.

The normative importance of the concept of overall freedom derives from a premise about the value of freedom that I mentioned earlier: that freedom has 'non-specific value', or value as such—in other words, that freedom has value independently of the value of being free to do one or another specific thing. This non-specific value of freedom can be either intrinsic or instrumental. It is perfectly consistent to affirm that freedom has only instrumental value—that freedom is only a means to an end—while also claiming that this instrumental value is of a non-specific kind (Carter 1999, Chap. 2). This will be so where the content of the end in question is unknown. One might know, for example, that freedom is the best means to economic or social progress, yet not know what this progress will consist in. One might affirm, indeed, that it is this very lack of knowledge that makes freedom the best means to progress, given that freedom allows us to experiment, to compare ideas, to make mistakes and to learn from them. In this case, our ignorance about the direction in which progress will take us makes it impossible for us to know which specific freedoms have value as a means to its realization. All we know is that freedom is a means to progress. Freedom is valuable as such, but only instrumentally valuable. This line or reasoning can apply to individuals as well as to aggregates of individuals, and from a purely prudential point of view rather than by reference to morally good ends. For example, an individual might see her own freedom as non-specifically instrumentally valuable in prudential terms because she is unable to predict her own future desires and beliefs.

Stoppino makes a very similar claim about power, implicitly interpreting A's power over B as having (prudential) non-specific instrumental value for A. According to Stoppino, A's power over B has instrumental value for A because it is a means to obtaining the conformity of B's conduct to A's preferences, which in turn is a means to the realization of A's ultimate goals. Now, in political life it might seem that such conformity becomes, for A, an end in itself, because A, as a political actor, typically attempts to achieve conformity not only 'here and now' (the conformity of some specific piece of behaviour of B) but also conformity that is 'generalized over space' (and therefore applies to a wide range of actors) and 'stabilized over time'. When conformity displays these two properties (of being generalized and stabilized), Stoppino calls it 'guaranteed conformity'. And the pursuit of guaranteed conformity is, for Stoppino, just what political activity consists in (Stoppino 2001b, Chap. 8). This is *not* to say, however, that political actors necessarily see power as *intrinsically* valuable, as if power in this generalized and stabilized sense were necessarily something that is pursued for its own sake. It is only to say that power is valuable *as such* for political actors, given its non-specific value as a means to the realization of those political actors' ultimate goals, whatever those goals might turn out to be. Power has specific instrumental value for A to the extent that it is instrumentally rational for A to pursue conformity 'here and now'; it has non-specific instrumental value for A to the extent that it is instrumentally rational for A to pursue conformity that is 'guaranteed as such' (Stoppino 2001a, p. 234)—to pursue it, one might say, *as if* it were an end. Thus, in the same way as the social freedom of B has non-specific value for B, the power of A has non-specific value for A.

In the light of this fact, it becomes interesting to ask how A's overall power is related to B's overall freedom. How far is it true that the growth of A's overall power over B (which, given the non-specific value A attaches to her power, increases (ceteris paribus) the subjective value of A's situation) implies a diminution of B's overall freedom to act (which, given the non-specific value B attaches to her freedom, decreases (ceteris paribus) the subjective value of B's situation)? In order to answer this question fully, we should need to be able to measure overall social power as well as overall social freedom, and that is not something that I feel warranted in assuming. One necessary step in the right direction, however, will consist in rendering explicit the effect on B's degree of overall freedom of each of the forms of power A might exercise over B.

4.2 Threats and Anticipations of Violent Sanctions

I shall begin by looking at the case of power that is exercised through the threat of violence. As I have argued elsewhere (Carter 1999, Chap. 8), the distinction between specific freedoms and overall freedom allows us to say that as well as limiting the freedom with which B acts, A's threat of violence limits B's freedom to act. When A threatens violence against B in order to induce B to do x, A does not remove B's freedom either to do x or not to do x. Nevertheless, A does typically reduce B's *degree of overall freedom* (to act).

To see this, we need to note that an agent's overall degree of freedom is not a function simply of how many members of a set of specific actions that agent is free to perform. In the first place, the sum of the courses of action one has available is not a sum of single actions, but a sum of various possible combinations of actions. I am probably free at this moment to shoot a policeman on Tuesday, free to shoot one on Wednesday, and free to shoot one on Thursday, but I am probably unfree to shoot three policemen (one on each of these days), given that I would probably be locked up after the first shooting. We need, then, to take into account not simply the possibilies of single actions (and the sum of these) but the *compossibility* of those actions for the agent: if P is free only to do x, y or z, while Q is free to do any combination of x, y and z, it is clear that Q is, ceteris paribus, the freer of the two. More generally, we should say that an agent's overall freedom is a function of the agent's set of sets of compossible actions. In the example just given, P has available the set of sets of actions $[\{x\}, \{y\}, \{z\}]$ while Q has available the set of sets of actions $[\{x\}, \{y\}, \{z\}, \{x, y\}, \{x, z\}, \{y, z\}, \{x, y, z\}]$. In the second place, we need to take into account, for each set of actions, not simply the availability or non-availability of that set, but the *probability* of that set being rendered impossible by some other agent. All judgements about freedom regard the possibility of actions that occur subsequent to the time of the freedom being predicated of the agent, and all such judgements are therefore most appropriately understood as probabilistic. And it would surely be grossly counterintuitive to describe as equally free, ceteris paribus, agent R, who is (at time *t*) 99 % certain to be prevented from performing a given set of actions, and agent S, who is (at time *t*) 1 % certain to be so prevented (Carter 1999, Sects. 7.5 and 8.4).

Bearing in mind these two factors of compossibility and probability, we can see that B's overall freedom should be understood as depending on the sum of all the sets of theoretically compossible actions for B, each one multiplied by the probability (between 0 and 1) of that set being rendered impossible by the actions of some agent, A (in the event of B attempting to perform that set of actions).[7] Given that we are talking of the prediction of the preclusion of a given set of actions (given certain conditions), and given that that prediction takes account of the probability of the preclusion, we may call the fundamental quantity determining B's level of overall freedom *B's overall degree of expected preclusion*.

It should already be clear at this point how B's overall degree of expected preclusion will, in the vast majority of cases, increase as a result of A's coercing B by threatening violent sanctions. The exercise of this form of power by these violent means generally implies, with a certain probability, that two or more actions that were compossible for B before the threat are now no longer compossible for B. Indeed, while A does not remove *any* specific freedom of B, B nevertheless suffers an increase in her overall degree of expected preclusion. Assume that A, in threatening B, does so with a minimum of determination and is minimally competent in carrying out the sanction. (These two requirements can be called the requirements of determination and competence, and we may call a threat that satisfies these requirements a 'true' threat.) In this case, at the moment at which the threat is issued (and indeed, even at the earlier moment at which A forms a resolute conditional disposition to impose the sanction (should B fail to comply)), A is *actually* (and with a certain probability) physically preventing B from performing at least one *set* of actions. A is actually precluding this set of actions to the extent that the counterfactual 'if B did *x*, A would do *y*' is true (where '*y*' is an action that prevents B from doing something). For an agent is

[7] What is the exact meaning of 'theoretical compossibility' in this context? This issue is problematic and has given rise to some debate in the literature. For Steiner (1994, Chap. 2), it means 'logically compossible'. In *A Measure of Freedom* I tentatively suggest that it might mean either 'logically compossible' or 'technologically compossible' or 'possible according to laws of nature' (Carter 1999, p. 173). For discussion, see van Hees (2000, pp. 131–133). Kramer (2003, Chap. 2) defines theoretical possibility, in this context, in terms of the agent's abilities, identifying freedom with ability and unfreedom with the prevention of that which the agent would otherwise be able to do. On this view, those actions the agent would be unable to perform even in the absence of prevention on the part of others, are classified as actions the agent is neither free nor unfree to perform: if I am unprevented from doing *x* but am nevertheless unable to do *x*, then I am neither free nor unfree to do *x*.

actually unfree to do something if it is true that, were that agent to attempt to do that thing, some other agent would intervene so as to render it impossible.

The truth of the counterfactual 'if B did *x*, A would do *y*' depends on A's dispositions to act, of which the communication of the threat is in fact only an indicator. Nevertheless, since this indicator is a fairly reliable one,[8] we can conclude that while the mere *threat* of a particular violent act is certainly distinct from the *actual* performance of that same violent act, the threat of that violent act is nevertheless generally accompanied by an *actual* increase in expected preclusion. When A truly threatens B with violence, A is (generally, and with a certain probability) *actually* preventing certain courses of action for B (i.e., certain sets of specific actions), regardless of whether B will comply with A's will or refuse so to comply.

To illustrate this point, let us return to the example of the bandit who says 'Your money or your life'. Assuming that the bandit is making a true threat (i.e., his threat satisfies the requirements of determination and competence), he is (at the time of the threat, and with a certain probability) physically preventing the respondent from holding on to her money *and* walking away, even though he is not preventing either the first or the second of these actions considered on its own. This follows from the truth (which is more or less probable at the time of the threat) of the counterfactual according to which, if the respondent chose to hold onto her money, the bandit would kill her. Similarly, in the case cited earlier of the oblique threat issued by Goebbels, the Nazi Government was (at the time of Goebbels' threat, and with a certain probability) physically preventing German citizens from writing certain things at time *t and* writing similar things at time $t + 1$ (*and* walking down the road unharmed at time $t + 2$, and so on), even though it was not preventing any of these actions considered in isolation from the rest. This reasoning shows how, even though there is no correlation between the threat of violence and *specific* unfreedoms, there is nevertheless a strong correlation between the threat of violence and *overall* unfreedom.

To be more precise about the difference between the effects of actual violence and the threat of violence, we need to note that each specific freedom is a member of *a certain number* of sets of actions that are compossible for the agent. Assuming, for simplicity, that the probability of prevention or non-prevention is always 100 %, the effect of A's *actual* violence is such that a certain specific action which was previously a member of at least one set of actions B was free to perform, is now no longer a member of *any* such set. The effect of A's *threat* of violence, on the other hand, is such that, while the number of sets of actions that B is free to perform diminishes, each of the specific actions that B was previously free to perform nevertheless remains a member of *at least one* of these sets. In other words, while actual violence removes *all* of the sets of which a given specific

[8] If it were not a fairly reliable indicator, then threats would fail as instruments of generalized and stabilized power. I return to this point at the end of the present subsection.

action is a member (from the list of sets the agent is free to perform), the threat of violence removes only some of these sets.

It is worth noting two consequences of this last point for the relation between degrees of violent coercive power and overall social freedom—consequences which largely reflect our pre-theoretical intuitions about that relation. First, the realization of an act of violence restricts freedom to a greater extent than the mere threat of that same act. Secondly, the more severe the act of violence threatened, the greater the power being wielded and the greater the reduction in the overall freedom of the agent who is subjected to that power.

Let us now continue to examine the case of violence (or of preclusion more generally), but in connection with another form of power identified by Stoppino: that of anticipated reactions. Here, although A does not issue a threat to B, the latter anticipates that were her own behaviour not to conform to A's interests a preclusive sanction would nevertheless be forthcoming. It should be clear that in such a case A limits B's freedom no less than where A issues a threat, for we have seen that the factor ultimately determining the restriction of B's overall freedom is the truth of the counterfactual 'if B did *x*, A would do *y*', and not the fact of A communicating this truth to B. The difference between A's threat of a sanction and B's correct anticipation of a sanction by A is not relevant, then, to the question of their effects on B's overall freedom. In Stoppino's formal classification, indeed, the essential difference between these two forms of power is that anticipated reactions represent a non-intentional (i.e., 'merely interested') exercise of power (See Table 1). And we have seen that the intentions of A, in precluding certain actions of B, are not relevant to questions about B's pure negative freedom (to act).

The correlation I have hypothesized between overall freedom and the threat or anticipation of preclusion is, in a sense, weaker than the correlation stipulated by those 'impure' negative theorists who simply define freedom as the absence not only of preclusion but also of punishability. The connection implied by my own analysis between these forms of power and overall freedom is not a logical, stipulative relation, but an empirical generalization. And, as in the case of all empirical generalizations, there will be exceptions to the rule. A first exception is where the requirements of determination or competence are not met: A either does not intend to carry out the threatened sanction (A is in fact bluffing) or is unable to do so (A overestimates A's own capacities), yet, since B is unaware of this, A's threat or B's anticipation of A's reaction still represents a successful exercise of power by A (i.e., B's choice still conforms to A's will in a way that it would not have done had B been fully informed). A second possibility is that the threatened or anticipated sanction would consist in A's inflicting harm on some third party, C, whom B cares about (hence the success of the threat), rather than on B herself.

These counterexamples are of limited relevance, however, for the study of political power relations and of their implications for political and social freedom (Carter 2008). As we have seen, in political life agents seek what Stoppino calls the 'guaranteed conformity' of the behaviour of others, and this implies conformity that is both *generalized* (over a large number of other agents) and *stabilized* (over time). The role of bluffs or incompetent threats in the pursuit of this guaranteed conformity cannot be anything but trivial, for it is clear that agents who fail to carry out sanctions fail to exercise generalized and stabilized coercive power. The counterexample of sanctions aimed at third parties is not answerable in the same way. Nevertheless, such sanctions are very rarely found in legal systems, and the reasons again have to do with the nature of political power relations. One such reason is that such sanctions are difficult to *generalize* as an effective instrument of power over many agents: different agents would react to them in more varied and unpredictable ways than they do to the threat of sanctions against themselves, and it is difficult to formulate general laws specifying the identities of the relevant third parties. But the most important reason is that coercive power exercised through the threat of sanctions against third parties would be difficult to *stabilize*, given the resentment and sense of injustice to which they would give rise. This resentment and sense of injustice would provoke a reaction on the part of the governed, and governments generally anticipate this reaction. This is itself an exercise of power by the governed over the government.[9]

4.3 Threats and Anticipations of Economic Sanctions

Let us now extend our analysis, within the forms of power consisting in coercion and anticipated reactions (understood as anticipated sanctions), beyond those cases where the relevant resources used by A are resources of violence (or more generally, resources permitting A directly to preclude certain act-combinations of B). Threatened or anticipated sanctions can also make use of economic or symbolic resources. In these cases, the application of the sanctions would not directly modify B's body or physical environment, but their conditional imposition by A nevertheless amounts to coercive power over B, as long as it is actually sufficient to induce B to modify her behaviour in A's interests. Examples of economic sanctions include fines imposed by the state, firings by employers, and industrial action on the part of unions. Examples of symbolic sanctions include stigmatization, exclusion from the community of the faithful, and eternal damnation.

Consider first the case of an economic sanction. A's firing B (where B is the employee), or A's going on strike (where B is the employer) brings with it a reduction in the economic resources available to B, which in turn would have constituted means by which B might have convinced other agents not to prevent B from performing certain actions. Economic sanctions imposed on B logically entail reductions in B's economic resources; the possession of economic resources logically entails the possession of economic *power*; and one's possession of economic

[9] For a more direct attempt to rebut this second counterexample, by showing that it fails to identify a threat that has no effect on B's set of sets of available options, see Kramer (2003, pp. 195–204).

power contingently (but nevertheless very strongly) affects one's degree of pure negative freedom. For example, when I buy an airline ticket I obtain, in exchange for a certain sum of money, a vast increase in the probability that I will not be prevented from boarding a certain aeroplane at a certain time. In the absence of this payment, I would very probably be prevented from boarding the aeroplane were I to attempt to do so (moreover, were I to succeed in boarding it on attempting to do so, I would very probably be punished afterwards). Therefore, if I do not possess the resources necessary to buy the airline ticket, my boarding the aeroplane (and my moving my body from Italy to the USA, and my visiting the Metropolitan Museum in New York, and so on) is something I am unfree to do. If I then acquire just enough resources needed to buy the ticket, I acquire that freedom to perform that action (and to visit the museum, and so on)—although it remains true that none of my sets of compossible actions contain the boarding of the plane without also containing the handing over of the money, and that none of my sets of compossible actions therefore contain the boarding of the plane, the handing over of the money, and some third action, *x*, the freedom to perform which would similarly depend on the handing over of the money. My boarding the plane will only become compossible with my doing *x* when I have doubled my money. And so on. (For an argument along these lines about the relation between freedom and money, see Cohen 2001).

It is therefore reasonable to say that in a market society characterized by well enforced rules of private property, there is a very strong causal link between a reduction in the market value of the resources at my disposal and an increase in my degree of expected preclusion. This is not to say, of course, that economic power is *essentially* the possibility of bringing it about that other people do not prevent one from doing certain things. But economic power does include that power among others. To possess economic power is to have the possibility of exercising (economic) coercion or remuneration; an exercise of coercion and remuneration is the bringing about of behaviour on the part of others; and that behaviour on the part of others often includes a series of door-openings. Moreover, it is enough for the agent to *possess* such power (without necessarily *exercising* it) in order to *possess* (with a certain probability) the set of sets of pure negative freedoms that would be brought into existence through those door-openings. After all, the relation we are examining here is that between A's power considered as an *exercise* concept and B's freedom considered as an *opportunity* concept—i.e., the effect of A's exercise of power on B's freedom to act. B *has* freedom (to act) as a result (*inter alia*) of B's *having* economic power (opportunity concept), given A's forbearance from *exercising* economic power over B.

4.4 Threats and Anticipations of Symbolic Sanctions

The limitation of pure negative freedom accompanying the threat or anticipation of *symbolic* sanctions is less immediately obvious. Nevertheless, its occurrence in political and social life is widespread and may be significant in terms of the degrees to which freedom is limited. Take, for example, the sanction consisting in exclusion from the community of the faithful. This sanction is likely to imply, indirectly and in the long term, the preclusion of a large number of options as a result of the future lack of collaboration (with B's endeavours) on the part of the faithful. A similar point will apply to most other cases of stigmatization and social exclusion.

This consideration does not, of course, apply to *all* symbolic sanctions. It does not apply, for example, to the sanction consisting in eternal damnation, *if* (and this may be a big 'if') that sanction is to be understood in the narrow sense of an event of disvalue that occurs only in the hereafter. It might be, that is, that a priest can exercise power over an individual by means of the threat of eternal damnation, even though no one in this world would have been any the wiser if, counterfactually, the individual had sinned and incurred eternal damnation. Such an individual does not incur a loss of freedom in *this* world as a result of the priest's exercise of power, for the requirement of competence has not been met. However, it does not seem to me counterintuitive, from the point of view of the theorist of negative freedom, to classify such a case as one of 'power without unfreedom', for in this example eternal damnation is not a punishment imposed by another agent or agents. The threat takes place in this world, but the threatened sanction (if it occurs) does not, and so does not involve the prevention of any actions. I shall call symbolic sanctions of this kind 'purely symbolic' sanctions. A purely symbolic sanction is one the realization of which would not be accompanied by any increase in preclusive behaviour either by the agent dispensing the sanction or by any third parties. Most symbolic sanctions, however, are not 'purely symbolic' in this sense.

One reason why the limitation of overall freedom is often less obvious where the threatened sanction is symbolic than where it is economic or violent, lies in the greater length of the causal chain of events linking A's intervention and the set of hypothetical actions (of third parties) which, at the moment of the threat, preclude certain acts or act-combinations of B. For example, a symbolic sanction imposed by A on B might bring it about that C imposes on B some economic harm, and only as a result of this that D (together with E, F …) physically prevent B from performing some act or act-combination (without preventing any specific actions). Presumably, it is correct to say that the length of the causal chain should influence our probability judgements about the likelihood of the actions of third parties that would prevent certain acts or act-combinations of B, and with these our judgements about B's degree of overall freedom at the time of the threat. Naturally, it is also possible that in the case of a symbolic sanction the causal chain is *shorter* than in the economic case. For example, it is possible that the fact of publicly labelling B as belonging to a certain race would straightforwardly induce C to act violently towards B.

It is certainly true that violent sanctions are more likely (than are symbolic sanctions) to be *disvalued* by B *because of* the increase in expected preclusion accompanying them. In threatening a violent sanction, A tends to be playing directly on B's desire not to have options closed off (along with other desires of B, such as that of avoiding pain). In threatening a symbolic sanction, on the other hand, A may only be playing on the intrinsic value B attaches to the symbol in question, or on its instrumental value for B in achieving other symbolic goods (such as recognition or self esteem). But this fact does not constitute an objection to my analysis. First, we should bear in mind that B's reason for disvaluing the threatened sanction (i.e., what makes the sanction count *as* a sanction, and therefore what makes A's intervention count as an exercise of *power*) is not in itself relevant to the question of whether and how far A is restricting B's *freedom*. For the desires of B, like the intentions of A, are not relevant to questions about B's freedom. Secondly, I would submit that the increases in expected preclusion generally accompanying a symbolic sanction *do* often contribute significantly to the disvalue B attaches to the sanction (and therefore to its counting *as* a sanction). B's stigmatization or social exclusion will no doubt have disvalue for B in terms of a reduction in self-esteem, but only in rare cases will its disvalue not be contributed to in some measure also by an accompanying non-trivial increase in B's degree of expected preclusion. Given this last fact, the connection between symbolic sanctions and restrictions of freedom is a less contingent one than might at first have been expected.

4.5 Direct and Indirect Restrictions of Freedom

Apart from the above-mentioned differences between different threats (or anticipations) in terms of the degree of restriction of B's overall freedom, we should also note a difference between threats (or anticipations) that involve A *directly* restricting B's freedom and threats (or anticipations) that involve A doing so *indirectly*. If A's power over B involves a *direct* restriction of B's freedom, this is because A is not only the threatener but also the agent of the counterfactual preventive actions that ultimately preclude certain acts or act-combinations of B. If A's power over B involves an *indirect* restriction of B's freedom, on the other hand, this is because those counterfactual preventive actions are actions of third parties (C, D, E …). In the case of a direct restriction of B's overall freedom, A's disposition to act (more precisely, A's conditional disposition to impose the sanction) is sufficient to determine that restriction. In the case of an indirect restriction of B's overall freedom, A's disposition is no longer sufficient to determine that restriction, which depends in addition on the conditional dispositions of C (D, E …).

In the case of violent threats (and anticipated reactions), A may be restricting B's overall freedom directly. Even violent threats, however, can be cases in which A is only restricting B's overall freedom indirectly. For example, A might threaten to order C (over whom A has power) to assault B physically. In the case of economic and symbolic threats (and anticipated reactions), on the other hand, A's restriction of B's overall freedom is never more than indirect. Where the threatened sanction is economic or symbolic, although the ultimate preclusion of certain acts or act-combinations of B would occur only if A imposed the sanction, that ultimate counterfactual preclusion is not itself brought about by A.

Now it might be objected that what I have called A's *indirect* restriction of B's freedom is not really a restriction of B's freedom at all. For in such a case, the counterfactual preventive actions of C (D, E …) are voluntary, at least in the minimal sense mentioned earlier, and often in the more demanding sense of being 'freely performed' (for it need not be the case that A, or indeed anyone else, is exercising power over C (D, E …) in this respect). Given this, it might seem that the most we can ever say, in the case of A's economic or symbolic power over B, is that C (D, E …) *would themselves* restrict B's freedom *were* A to carry out the sanction. And since A *does not* actually carry out the sanction (for we are assuming that B complies and A's exercise of power is therefore successful), no such restriction of B's freedom actually occurs.

This objection assumes that in order for us to *impute to A* a restriction of B's freedom that depends on the hypothetical actions of C (D, E …), A must somehow *cause* those actions of C (D, E …), in such a way as to deny the free agency of C (D, E …). It does not seem to me, however, that A must be (counterfactually) the cause of the preventive actions of C (D, E …) in order to be one of the actual causes of an increase in B's overall degree of expected preclusion—an increase that in fact takes place when A forms the resolute disposition to carry out the sanction should B not comply. Although the actions of C (D, E …) are voluntary, it remains true that these actions would take place *if and only if A imposed the sanction*. The increase in B's overall degree of expected preclusion therefore depends on A, and this fact is sufficient to motivate the claim that the disposition of A contributes to that increase.

This point can be argued more technically by assuming the analysis of 'sources of unfreedom' recently presented by Kramer (2003, Chap. 4). A human action is a source of an agent's unfreedom to do *x* if it contributes causally to the state of affairs in which it is impossible for the agent to do *x*. What is the relevant meaning of 'contributes causally' in this context? The answer is that an event or state of affairs X contributes causally to the occurrence of another event or state of affairs Y if and only if X passes the so-called NESS test—i.e., the test of whether X is a 'necessary element of a sufficient set' of conditions for Y.[10] To pass this test, X must be a member of a set of minimally sufficient conditions for Y. The fact that the set is 'minimally' sufficient implies the necessity of each and every member of the set for the realization of Y, in the sense that if any one such member were not to be realized, Y would not be realized either. Kramer rightly sees this causal criterion as implied by any attempt to distinguish clearly and plausibly between obstacles that are sources of *social* unfreedom (because they are contributed to by human agency), and other obstacles that are instead to be classed as being of purely natural origin, or else as self-inflicted. Moreover, and more relevantly for our purposes, the same causal criterion also serves to single out *which* of various human agents are contributing to a given social obstacle, where more than one such agent appears to be doing so. In the example just discussed, the actual

[10] A useful account of this test is given in Braham and Holler (2009).

(conditional) disposition of A, to impose a sanction on B should B not comply, is certainly a necessary member of a set of minimally sufficient conditions for an increase in B's degree of expected preclusion to occur, no less than is the actual (conditional) disposition of C (D, E …) to act in certain ways should A impose the sanction. The set of dispositions of A and C (D, E …) is a set of minimally sufficient conditions for B's increase in expected preclusion. And since that set of dispositions is *actually* realized, so is B's increase in expected preclusion.

It should be noted that the kind of causation Kramer is talking of is *physical causation*, which takes into account, as variables, *all* the physical events and states of affairs that are relevant to determining the range of options physically available to B—among which, for example, the presence of oxygen in the air. Stoppino, on the other hand, talks of power as a causal relation between the conduct of A and that of B. When Stoppino talks of causation, then, he has in mind *social* causation, which assumes B's physical environment (as well as B's utility functions) to be fixed, and takes account, as variables, only of the conduct of A, C, D, E …. This allows Stoppino to state that, when A exercises power over B, A's conduct is in itself a *sufficient cause* of B's conduct (Stoppino 2001a, pp. 8–11). (This would certainly be false if among the necessary conditions of B's conduct we were to include the oxygen surrounding B, B's utility function, and so on). But this difference—between physical causation and social causation—is not relevant to the thesis I am defending, according to which A limits B's freedom when A (determinedly and competently) threatens B with an economic or symbolic sanction that would induce C (D, E …) to prevent B from performing certain sets of actions. For the same conclusion about B's unfreedom that follows from Kramer's causal criterion also follows from Stoppino's. Indeed, when I stated earlier (in defining direct versus indirect restrictions of freedom), that in the case of A's restricting B's freedom *directly*, A's disposition to act is *sufficient* to determine that restriction, I was assuming the *social* concept of causation. A's disposition is sufficient to determine that direct restriction of freedom if (and only if) we treat as variables only the behaviours and dispositions of A, C, D, E …. Where A limits B's freedom *indirectly*, on the other hand, what matters is that A's disposition be a necessary element of a set of behaviours and dispositions that is minimally sufficient to determine the restriction of B's freedom. And this *is* a feature of A's disposition in the case of indirect restrictions of freedom, on Stoppino's causal criterion no less than on Kramer's.

4.6 The Counterfactuals of Indirect Restrictions of Freedom

It is worth pausing at this stage to make explicit the nature of the counterfactuals in play in the case of the indirect restrictions of freedom accompanying coercive power. The first thing to note is that, while A's direct restriction of B's freedom (accompanying A's coercive power) is entailed by the truth of a single counterfactual (if B did *x*, A would do *y*), A's indirect restriction of B's freedom (accompanying A's coercive power) depends on the truth of the conjunction of *several* counterfactuals ordered in a chain, such that with each successive link in the chain we progressively distance ourselves from the actual world (in which B conforms to A's will by refraining from doing *x*).

In the case of any one concrete example of such an indirect restriction of freedom, the composition of the relevant chain of counterfactuals can be built up from one or more chain segments. I suggest we think of these chain segments as coming in two standard forms. The first and simpler form is a segment made up of two elements, which consist in the following two counterfactuals:

– if B did *x*, A would do *y*;
– if A did *y*, C (D, E…) would prevent B from doing *a*.

This chain segment renders explicit the sense in which A indirectly restricts B's overall freedom in the simple example where A threatens to stigmatize B, and as a result of that stigmatization C (D, E …) would act violently towards B.

The second form of chain segment is made up of three elements, where the ultimate preventive actions of C (D, E …) depend on further actions or omissions by B (made inevitable by A's sanction):

– if B did *x*, A would do *y*;
– if A did *y*, B would be incapable of doing *r*;
– if B did not-*r*, C (D, E …) would prevent B from doing *a*.

This chain segment renders explicit the sense in which A indirectly restricts B's overall freedom in the case where A exercises coercive economic power over B. For example, where A threatens to fire B (and thus to harm B economically), A's carrying out of the sanction would diminish B's capacity to continue to exercise remunerative power over C (D, E …) in such a way as to bring it about that C (D, E …) do not prevent B from doing certain things. (Typically, in this chain segment '*r*' is a *set* of remunerative acts by B and '*a*' is some further *act-combination* for B.) The actual preventive dispositions of C (D, E …) are here conditional in two senses: first, they are directly conditional on B failing to make certain payments; secondly, because B's failing to make certain payments is conditional on A's economic sanction, they are indirectly conditional on that economic sanction. The conditionality is two-fold because the chain segment has three elements instead of two.

Where the chain of counterfactuals needed to render explicit an indirect restriction of freedom is longer still, it can be reconstructed by assembling instances of the two forms of chain segments just set out. We might, of course, need a very long chain made up of very many segments. But to illustrate, take the relatively simple case of A threatening to impose on B a symbolic sanction (for example, exclusion from the community of the faithful) which would result in a third party imposing an economic cost on B (for example, in B losing her job). To account for A's indirect restriction of B's freedom in this case, we shall need one of each of the two forms of segment described above—a two-piece segment followed by a three-piece segment. In the following list of counterfactuals, the second counterfactual constitutes both the final link in the first segment and the first link in the second segment:

– if B did x, A would do y;
– if A did y, C would do w;
– if C did w, B would be incapable of doing r;
– if B did not-r, D (E, F …) would prevent B from doing a.

It should be emphasized once more that in none of these examples of coercion is an actual chain of events being described. *Ex hypothesi*, A succeeds in exercising power over B, so that in the actual world B does not do x and the chain is cut off at the first link. A's restriction of B's overall freedom does not depend on the realization of any of the events referred to in the consequents of the above conditionals, but only on the truth of the conditionals themselves.

4.7 Remuneration Versus Coercion

Stoppino classes coercion (the threat of sanctions, both violent and non-violent) and anticipated reactions (the anticipation of sanctions, both violent and non-violent) as forms of power that work through direct interventions on B's 'available alternatives'. ('Available alternatives' should be understood here not in the objective sense of unprevented courses of action—the sense that the theorist of pure negative freedom would have in mind in using the term—but in the subjective sense of 'the various courses of action that B takes into consideration' (Stoppino 2007). They are B's available courses of action weighted according to their degree of eligibility in B's eyes).

A third form of power that fits into this same category is that of 'remuneration'—that is, the promise on the part of A to reward B should B perform a certain action. Moreover, although the anticipated reactions discussed so far have all been anticipated sanctions, we should not forget that anticipated reactions can also be anticipated rewards.

It follows that, even leaving aside the exceptions mentioned above, not all of the forms of power whose 'target of intervention' consists in B's 'available alternatives' are forms of power that result in restrictions of B's overall freedom. For where A exercises power over B by promising a reward, A is generally *increasing* B's set of sets of compossible actions, and thus B's overall freedom. This result reflects the pre-theoretical intuition of most liberals, assuming a broad sense of 'liberal'. Indeed, the implication that A could restrict B's freedom by actually *increasing* B's set of available actions (and vice versa) ought to sound alarm bells in the mind of any liberal political theorist.

There can of course be offers that B *accepts unfreely*, given a certain definition of acting freely. This might be said, for example, of the offer made in the film *Indecent Proposal*, where a millionaire offers an enormous sum of money to a married couple on condition that the woman spends the night with him. Even in such cases, however, A is generally increasing B's overall pure negative freedom (to act), even if B complies unfreely.

It must again be emphasized, however, that the claim that remuneration increases B's freedom, like the claim that coercion limits it, is a contingent one. There are offers that are not accompanied by any change in B's degree of overall freedom, for reasons that exactly mirror those mentioned in the case of threats: A might not be determined or competent to keep the promise, or the benefit offered by A might not in any case be such as to reduce, directly or indirectly, B's degree of expected preclusion, consisting instead in some other event that B values, such as a benefit to a third party about whom B cares.

As a final observation on coercion and remuneration, we should note the possibility not only of A's leaving B's freedom unaffected (in the case both of A's threatening a sanction and of A's promising a reward), but also the possibility of A's increasing B's overall freedom through a threat and of A's reducing it through an offer. Coercion counts as such if the behaviour threatened by A is seen by B as contrary to B's own interests. And it is always possible for B to judge an increase in her own degree of expected preclusion to be in her own interests. This will be so where B desires to have her own choices restricted for her own good. In this case, A's freedom-restricting intervention will count not as a sanction, but as a reward. Similarly, A's posing the condition that if B does *x*, A will cease preventing a certain course of action, will in this case count as coercion. It is important for my general thesis about the correlation between A's coercion and B's unfreedom that such preferences on the part of B be exceptional. Most of us find this a reasonable assumption, and I think that the explanation lies in our assumption that the rationality of preferring more freedom to less tends to be overturned only in limited circumstances. Two such circumstances are worth mentioning here. First, B might prefer being prevented from doing certain things because B wishes to be protected against her own weak-willed desires. Such a preference, however, generally occurs only with respect to a very limited number of pursuits. Secondly, B might find that an abundance of available alternatives negatively affects her capacity to make a rational choice, within given time-constraints, between the specific alternatives open to her (Dworkin 1988, pp. 66–67). The most widely cited example is that of a choice between products in a supermarket: it might be rational for the agent to prefer having a choice between six decent brands of toothpaste to having a choice between sixty-six. Nevertheless, this preference of B only kicks in when her level of freedom is above a certain threshold. Moreover, the preference is only likely to apply to certain kinds of freedom. The claim that such preferences exist is plausible when applied to those freedoms that necessitate the exercise of our faculty of rational choice within strict temporal constraints, such as the choice of a toothpaste in a supermarket. But it is much more difficult to find similar examples in areas where a time-constraint is neither objectively present nor self-imposed. In considering the traditional liberal freedoms of worship, of association or of movement, for example, we do not tend to think that there is a threshold above which increases in the number of options take on a negative value for the agent.

4.8 Informational and Psychological Manipulation

In Stoppino's classification, there are three other categories of power (each comprising one or more particular forms of power) where the target of A's intervention is something other than B's 'available alternatives' (see Table 1). These three other possible targets of A's intervention are as follows.

First, A might intervene on B's 'factual knowledge and value beliefs'. In this first category we find the forms of power that Stoppino calls 'informational manipulation', 'persuasion', and 'imitation'. For example, A might induce B to engage in certain forms of behaviour by indoctrinating B ideologically or by convincing B of the validity of certain beliefs by means of rational argument.

Secondly, A might intervene on B's 'unconscious psychological processes'. In this category we find the form of power called 'psychological manipulation'. Examples cited by Stoppino are subliminal advertising and brainwashing.

Thirdly, A might intervene on B's 'social environment', either by modifying the dispositions-to-act of third parties, so as to induce them to change the behaviour of B in A's interests, or by modifying the distribution of resources, so as to modify B's preferences in line with A's interests. An example Stoppino gives of a modification of third parties' dispositions-to-act is that of a couple that finds itself unable directly to influence the behaviour of their rebellious son. As a result, the couple adopts the alternative strategy of somehow convincing a third party (friends or family) to change the son's behaviour in line with the couple's wishes. An example Stoppino gives of a modification of the distribution of resources is that of the acquisition by A (by means of purchases from C, D, E …) of a monopoly of a certain kind of resource that B needs. As a result of this acquisition, B modifies her behaviour in line with A's interests in order to guarantee the availability of this resource. Power that is exercised through an intervention on B's social environment can be called 'indirect' power, since the causal relation between A's conduct and B's will is mediated by an intervention on some outside factor that in turn modifies B's beliefs and desires and/or B's perception of her available alternatives.

In this third and last category—the category of power that works through an intervention on B's social environment—we find the forms of power called 'situational manipulation' and 'conditioning' (where the latter can be either intentional or merely interested). Manipulation being a 'hidden' form of power (where A keeps B unaware of the power relation or its nature), situational manipulation represents the 'hidden' version of power that is exercised through an intervention on B's social environment, whereas conditioning represents the 'open' version (where A does not hide the power relation or its nature). Because of its hidden nature, the occurrence of situational manipulation tends to be limited to small groups of agents. (These, however, might be very powerful groups, such as a government executive, in which case the consequences of situational manipulation can still be far-reaching). Conditioning, on the other hand, can play a more direct role in the successful implementation of public policies.

Of these three additional targets of A's intervention (B's factual knowledge and value beliefs, B's unconscious psychological processes and B's social environment), the one that most obviously identifies a category of power where A restricts B's freedom is the third: the intervention on B's social environment. I shall turn to this category in the next sub-section.

As I stated at the outset, there are some forms of power, such as rational persuasion and imitation, that no theorist of negative freedom would see as a restriction of B's freedom. More controversial is the question of whether A can restrict B's freedom by withholding information from B, or by the strongest forms of psychological manipulation. For the theorist of pure negative freedom, this question will turn on whether such conduct on the part of A really makes certain actions of B *impossible* (Carter 1999, p. 206; Kramer 2003, pp. 82–83, 255–271).

On the basis of Stoppino's classification, it is reasonable to say that there are some forms of informational or psychological manipulation by means of which A restricts B's pure negative freedom, and others by means of which A does not do so. As far as information is concerned, it is important to distinguish between 'knowing how' to do something and 'knowing that' something is the case. B's freedom to do *x* at time *t* will be removed by A if A withholds from B information without which B cannot possibly *know how* to do *x*. In an example given by Kramer (2003, pp. 82–83), B is locked in a room and told that the door will open only if she punches 200 digits on a keyboard in exactly the right order. In this case, B's ignorance of the code makes her unfree to exit within a certain time limit (or more precisely, very probably unfree to do so). Consider, on the other hand, a case in which A leads B to believe that she has been locked in a room by A when in fact the key has not been turned. In this case, what B lacks is not knowledge about *how* to exit, but knowledge *that* she is unprevented from exiting. This last kind of ignorance is not a source of pure negative unfreedom. For consider the test we must apply in order to see whether (or better, with what probability) others have made it impossible for B to do *x*. This test consists in asking, 'Were B to try her best to do *x*, would B fail to do *x* as a result of the actions of others?'. In the case of B's ignorance of the code needed to exit the room, the answer to this question is 'yes' (or better, 'very probably'), whereas in the case of B's ignorance of the door being unlocked, the answer is 'no' (or better, 'very probably not, and in any case with no higher a probability than had B known of the door being unlocked). It seems reasonable to say, then, that being prevented from 'knowing how' to avail oneself of an option is a source of unfreedom, whereas being prevented only from 'knowing that' one has the option is not a source of unfreedom. A more realistic and politically relevant example of people being rendered unfree to do certain things through a denial of 'know-how' would be where a government denies to certain classes of people—for example, to women or to certain races—the possibility of frequenting certain university courses. If a 20 year-old is prevented from studying medicine for the next 10 years, she is, at the moment of that prevention, rendered unfree to carry out a certain medical operation at the age of 30.

Many instances of 'knowing that' something is the case will be instrumental to 'knowing how' to perform a particular action, and thus to increasing the probability of one's succeeding in performing that action should one try. For example, it may be that in order to take the train from *a* to *b*, I need to 'know that' the train to *b* is the one whose eventual destination is *c*. Knowing that this is the case is an example of 'knowing how' to take the train to *b*, even though it is formulated in the language of 'knowing that'. Similarly, our medical student will know how to carry out a heart operation successfully only if she knows that the heart is structured in a certain way, and the person locked in the room in Kramer's example will know how to exit only if she knows that that code consists in a certain sequence of numbers. When a statement about 'knowing that' can be reformulated as a statement about 'knowing how' to perform a certain action, then the prevention of the knowledge referred to can constitute a restriction of one's freedom. The claim that one 'knows that' *one has a certain option*, however, cannot be reformulated as the claim that one 'knows how' to avail oneself *of that option*. It is for this reason that ignorance about the existence of available options does not make one unfree to avail oneself of those same options.

But while it is true that the withholding of know-how restricts the freedom of those who are thereby kept in the dark, it is not clear that this activity qualifies as a form of *power*. The reason for this is that A's withholding of information about how to do *x* simply renders B unfree to do *x*. Where A exercises *power* over B, on the other hand, A brings it about that B does not-*x while nevertheless leaving B free to do x*. The withholding of know-how, then, is a case of preclusion, not of power. Informational manipulation will count as an exercise of power only when it is a withholding of factual information the effectiveness of which (in making B refrain from doing *x*) does not depend on its usefulness to B in understanding how to do *x*. Lying, suppressing information, and providing excessive information—all examples provided by Stoppino—*may* still count as power on this view, but in many concrete instances they will not.

Since it is reasonable to call manipulation a form of power, I shall call the cases just cited, in which A restricts B's freedom by withholding information, cases of 'informational preclusion' rather than of 'manipulation'. A's use of information to modify B's behaviour can be an exercise of power or a case of preclusion. Where it is an exercise of power, it will not constitute a restriction of freedom. Where it does constitute a restriction of freedom, on the other hand, it will not qualify as power (being informational preclusion). The fundamental reason for this is that *if* A restricts B's freedom by means of informational preclusion, A precludes *certain specific freedoms of B*, whereas when A exercises *power* over B, A necessarily precludes at most certain *act-combinations* of B.

The case of psychological manipulation is identical to that of informational manipulation in this respect. There are clearly extreme cases of intervention on B's unconscious psychological processes—for example, brainwashing as described by Stoppino—in which A renders B's performance of certain actions impossible by making certain psychological processes impossible for B. However, in these extreme cases what happens is that A precludes certain actions of B. A does not leave B free to do otherwise, and as a consequence A cannot be said to be exercising power over B. Therefore, if Stoppino is right to assume that in cases of power B is necessarily free to do otherwise, then he is wrong to class brainwashing as an exercise of power. This is a form of psychological preclusion, not of psychological manipulation. Less extreme kinds of intervention on B's unconscious psychological processes, on the other hand, do qualify as power, but they do not restrict B's freedom. For they neither remove any specific freedoms of B nor (in themselves) reduce B's level of overall freedom. Thus, theorists of negative freedom are not amenable to the suggestion that advertising or emotive religious or political propaganda (the use of symbols such as flags or prayers or anthems) renders people unfree to perform any specific actions or act-combinations. These latter kinds of intervention—advertising or emotive propaganda—affect people's inclinations, but it is essential to any negative conception of freedom that one make a clear distinction between being inclined not to do *x* and being unfree to do *x*.

It seems to me that Stoppino was led to classify brainwashing and the withholding of know-how as examples of power because he lacked the category of (intentional and interested) *preclusive* behaviour (which includes, but is not limited to, violence). Faced with the choice of classifying them either as power or as violence, it will have seemed more natural to place them in the former category. After all, neither can be easily qualified as an intervention on the agent's body or immediate physical environment (although brainwashing tends to be accompanied by such an intervention). We have seen, however, that it is difficult on reflection to justify classing them as examples of power. Instead, they should be seen as lying outside either category, but within the wider category of intentional and interested preclusion.

It should be noted that in cases of informational or psychological preclusion, A restricts B's freedom *directly*. A can, of course, bring it about that third parties engage in similar acts towards B, withholding know-how from B or engaging in brainwashing or hypnosis. In this case, A is restricting B's freedom indirectly. Where A does so, however, A is exercising power by intervening on B's social environment (i.e., on third parties' dispositions to act), and the form of power is therefore situational manipulation or conditioning. While the restrictions of freedom involved in informational and psychological preclusion are always direct, the restrictions of freedom involved in situational manipulation and conditioning are always indirect. (The only target of intervention that admits cases both of direct and indirect restrictions of B's freedom is that of B's 'available alternatives'. Coercion and anticipated reactions can involve A *in*directly restricting B's freedom without actually engaging in conditioning, because they do not involve A *actually* modifying third parties' dispositions-to-act.)

Finally, we should note that, like interventions on B's available alternatives, interventions on B's factual knowledge or unconscious psychological processes can involve increases in B's overall freedom as well as decreases (one might call these cases of informational or psychological 'enablement', as opposed to informational or psychological preclusion). The positive counterpart of A's depriving B of knowledge about how to avail herself of certain options is, clearly enough, A's supplying B with that knowledge. The positive counterpart of brainwashing is probably psychotherapy. (Stoppino mentions this as a rare example of open (i.e., non-manipulative) power that nevertheless has unconscious psychological processes as its target of intervention.) For example, if brainwashing can make B unfree to do *x* by inducing in B a particular phobia, and psychotherapy can remove that phobia, then A's acting as B's psychotherapist can result in A's directly increasing B's freedom. Even assuming B's consequent behaviour to conform to A's interests, however, A's increasing B's freedom through psychotherapeutic activity is no more an exercise of power by A over B than is A's restricting B's freedom through brainwashing.

4.9 Situational Manipulation and Conditioning

As I have suggested, B's pure negative freedom is certainly restricted by A in many cases in which A exercises power over B by intervening on B's *social environment* (i.e., cases of situational manipulation and conditioning). The way in which this comes about is particularly clear where A's intervention on B's environment is an intervention on third parties' dispositions-to-act. Such an intervention will constitute a restriction of B's overall freedom whenever the conduct of C (D, E …) produced by that of A is itself a restriction of B's freedom in one of the ways already discussed. For example, if A modifies B's behaviour (in A's interests) by persuading C to threaten B with an economic sanction, A contributes to the resulting reduction in B's freedom because, as in the earlier cases examined, A's intervention, no less than C's, is a necessary element of a set of minimally sufficient conditions for B's suffering an increase in her degree of expected preclusion. If, on the other hand, A modifies B's behaviour by *coercing* C into to *persuading* B, A may thereby be reducing C's overall freedom, but A is not thereby reducing B's overall freedom.

It should be noted that in these cases of power (both where A restricts B's freedom and where A does not), the conduct of C is not necessarily in C's interests, and therefore does not necessarily constitute an exercise of power by C. Nevertheless, the behaviour of C will always be *equivalent* to the exercise of one of the more direct forms of power (forms of power where the target of intervention is B's available alternatives or B's factual knowledge and value beliefs or B's unconscious psychological processes), where by its being 'equivalent' I mean that it consists in the same physical behaviour on the part of C, even though the behaviour C induces in B might not be one that conforms to C's interests. For example, a politician (A) might coerce an employer (C) into threatening to fire his employee (B) unless the employee gives his support to the politician in an election, even though the politician's being elected is not in the employer's interests. In this case, the politician is exercising power over the employee (here, via an exercise of power over the employer), but the employer's threat is not itself an exercise of power over the employee. The employer is instead engaging in behaviour that I am calling 'equivalent' to an exercise of (economic, coercive) power, as well as being a necessary link in A's indirect power over B.

Stoppino mentions the redistribution of resources as a method of situational manipulation or conditioning, in addition to the modification of third parties' dispositions to act. It is not clear, however, that A's effecting a redistribution of resources represents a genuinely distinct from of intervention on B's social environment—an intervention that is somehow an *alternative* method, for A, to that of effecting a change in third parties' dispositions-to-act. After all, *every* restriction of a person's freedom depends ultimately on the actions of others and thus on their dispositions to act. This point leads me to doubt the status of 'redistribution of resources' as an independent way of exercising situational manipulation or conditioning. It seems to me, indeed, that this type of intervention can always be categorized in one of the following two ways.

First, it might be categorized as a redistribution of resources that *brings about* a subsequent modification of *third parties' dispositions-to-act*. An example might be where A redistributes a certain resource that B needs from one third party to another—i.e., transferring the property rights in that resource from D to C—with the consequence that B's behaviour is modified in line with C's interests (which, unlike D's interests, happen to coincide with those of A). Here, the change in B's behaviour takes place because, thanks to the redistribution of resources in question, C is given the opportunity to prevent B from performing certain sets of actions (in the way already illustrated in connection with the threat and anticipation of economic sanctions), and then develops the disposition to do so should B not conform to her interests. This is certainly a case in which A exercises power (indirectly) over B. However, it is an exercise of power that works, ultimately, by means of a modification of *third parties' dispositions-to-act*. In this case, then, we do not seem to be justified in calling the redistribution of resources an *alternative* method with respect to the method of modifying third parties' dispositions to act. Rather, such redistribution is just one of the *means* by which A might conceivably bring about a change in third parties' dispositions-to-act.

A useful clarificatory example of a redistribution of resources (resulting in a modification of third parties' dispositions to act) is that of an *actual economic sanction* imposed by A on B. This actual sanction is a sign that A has attempted and failed to *coerce* B into acting in a certain way. Despite indicating a failure of A to coerce B, however, the sanction may also turn out to be a means by which A, intentionally or unintentionally, *conditions* B, given the expected behaviour of C (D, E …) consequent upon A's sanction, as already illustrated in the previous analysis of economic coercion—except that in the case of the *actual* economic sanction, C's (D's, E's …) behaviour is *actual* (because consequent upon an actual sanction by A), as opposed to *counterfactual* (because consequent upon a counterfactual sanction by A). In the case of the actual economic sanction, A redistributes resources from B to C (D, E …) and in so doing modifies C's (D's, E's …) *dispositions to act* towards B.

The second way of categorizing a change in the distribution of resources is as a redistribution of resources *without* any consequent modification in third parties' dispositions-to-act. In such cases, however, the power exercised by A over B should not be classed as situational manipulation or conditioning, but as one of the direct forms of power previously discussed. Here too, then, A's redistribution of resources

Table 2 The relation between social power and overall freedom

target of intervention / effect on overall freedom	social environment	unconscious psychological processes	factual knowledge and value beliefs	available alternatives
direct restriction				[h] *coercion; anticipated reaction* (true threat or correct anticipation of physical removal of options, i.e. of violence)
indirect restriction	[a] *situational manipulation*; *conditioning* (intervention inducing in third parties behaviour identical or equivalent to [h] or [i])			[i] *coercion; anticipated reaction* (true threat or correct anticipation of economic or symbolic sanction)
no effect	[b] *situational manipulation*; *conditioning* (intervention inducing in third parties behaviour identical or equivalent to [d], [e], [f], [g], [j] or [k])	[d] *psychological manipulation* (including use of emotive symbols, but not including brainwashing or hypnosis)	[e] *informational manipulation* (including withholding of information about available options, but not including withholding of know-how) [f] *persuasion* [g] *imitation*	[j] *coercion; anticipated reaction* (false threat or incorrect anticipation; true/false threat or correct/incorrect anticipation of 'purely symbolic' sanction or of harm only to third party) [k] *remuneration; anticipated reaction* (false promise or incorrect anticipation; true/false promise or correct/incorrect anticipation of 'purely symbolic' benefit or of benefit only to third party)
indirect increase	[c] *situational manipulation*; *conditioning* (intervention inducing in third parties behaviour identical or equivalent to [l], [m])			[l] *remuneration; anticipated reaction* (true promise or correct anticipation of economic or symbolic benefit)
direct increase				[m] *remuneration; anticipated reaction* (true promise or correct anticipation of removal of humanly imposed physical constraints)

fails to qualify as an independent form of situational manipulation or conditioning—in this case, because it fails to qualify as a form of situational manipulation or conditioning. Imagine, for example, that A succeeds in acquiring a monopoly over a certain kind of resource that B needs (this is Stoppino's example). Here, A redistributes resources (by acquiring them from C (D, E …)) but does not rely on any modification of third parties' dispositions-to-act in order to modify the behaviour of

B in conformity with A's interests. Nevertheless, in such a case it seems correct to classify the power exercised by A over B not as situational manipulation or conditioning, but as the direct threat (by A) or anticipation (by B) of a sanction—that is, as coercion or anticipated reaction. This exercise of power depends on A's disposition to impose economic sanctions on B should B engage in certain forms of behaviour—sanctions which, before acquiring the monopoly in question, A was unable to impose, and which A is enabled to impose as a result of the monopoly. Therefore, the exercise of power that takes place in this example is not itself exemplified by the acquisition of the monopoly (the redistribution *without* a consequent modification in third parties' dispositions-to-act). Rather, the acquisition of the monopoly is *previous* to A's exercise of *coercive* power (or B's anticipation of A's reaction). The acquisition is *preparatory* (be it intentionally or unintentionally) to A's exercise of coercive power over B (or B's anticipating A's reaction), since that exercise of coercive power requires *further* conduct on the part of A (or anticipation on the part of B). The acquisition itself is an act that creates *potential* power (an act that *gives* A power), by supplying A with new resources and hence new opportunities for imposing sanctions.

(In his discussion of situational manipulation, Stoppino says that, by secretly acquiring a monopoly of a given resource, A can 'just as secretly dictate his demands' on B (Stoppino 2007). To the extent that this is so, however, it suggests, *pace* Stoppino, that there are forms of coercion and remuneration that can be 'hidden' in Stoppino's sense, rather than that the secret acquisition of a monopoly is itself a separate form of hidden power working through an intervention on B's social environment. A's power is exercised through the secret dictation of his demands, which happens to follow his secret acquisition of the monopoly.)

Thus, although my analysis of the freedom-power relation has generally taken Stoppino's formal classification as given, I am nevertheless moved, in the light of that same analysis, to challenge two aspects of that classification. First, as we saw in the previous subsection, a number of cases that Stoppino would classify as ones of informational or psychological manipulation should not, after all, be classed as examples of power—even though, as we also saw, our very reason for not classing them as examples of power is also a reason for classing them as restrictions of freedom. Secondly, the kind of power that Stoppino calls situational manipulation or conditioning operating by means of a redistribution of resources (*rather than* by means of a modification of third parties' dispositions-to-act) is not, in reality, a separate form of power, but is more properly classed as the *creation of potential power*.

5 Conclusion

We have seen that there is a complex relation between A's social power and B's overall negative freedom, depending both on the form and the substance of the power relation. This complex relation is set out schematically in Table 2.

In line with the criticism of Stoppino presented at the end of Sect. 4.9, I have omitted from Table 2 those cases of situational manipulation or conditioning that Stoppino would class as operating by means of a redistribution of resources. I have also omitted cases of psychological or informational preclusion (in line with the criticism presented in Sect. 4.8), given that Table 2 concerns the relation between B's overall freedom and A's *power*, in the strict sense assumed by Stoppino. It should not be forgotten, however, that psychological and informational preclusion nevertheless represent restrictions of freedom, even though (and in a sense, because) they lie outside Table 2. These exceptions aside, Table 2 reproduces all the forms of power identified in Stoppino's classification (see Table 1), and similarly groups them according to the relevant target of intervention. Table 2 also contains two minor simplifications: first, in order to avoid overcrowding, the table omits cases of coercion that increase overall freedom and cases of remuneration that reduce it; secondly, in the case of threats (or offers, or anticipations) that are 'true' (i.e., that satisfy the requirements of determination and competence) but nevertheless have no effect on B's overall freedom, the table refers only to cases of sanctions (or benefits) affecting a third party.

Overall, the above analysis seems to me to provide a plausible account of the relation between social power and negative freedom. It shows how negative freedom is restricted not only by the most obvious relations of preclusion—in particular, violent relations of preclusion—but also by the less evident forms of preclusion that accompany a number of different forms of power, including coercion, anticipated reactions and many instances of situational manipulation and conditioning, as well as by informational and manipulative forms of preclusion. Thus, it is misguided to depict the pure negative conception of freedom as entailing a particularly 'narrow' or 'restrictive' view of the relation between power and freedom, as if A only limited B's freedom through violence or the physical prevention of specific actions. This, despite the fact that the pure negative conception does indeed entail that, ultimately, freedom is restricted only through the social preclusion of acts or act-combinations.

I would suggest, further, that the above analysis gains appeal from the fact that it lays down the basis for some potentially fruitful interaction between political scientists and normative political theorists. Any adequate normative political theory endorsing the aim of limiting, controlling or distributing certain forms of social power in certain ways must give a plausible account of the reasons for pursuing such an aim. It must ground that aim in a normative sense, by referring to the values that the control, limitation or distribution of power will ultimately promote. An important value commonly cited by liberal and republican theorists as a justification for the limitation of political power is the value of freedom. Such theorists believe that a measure of freedom, or equal freedom, or maximal equal freedom, is owed to individuals as a matter of right—either because our moral obligations include a fundamental obligation to respect other moral agents as such, or because they believe that we are obliged to respect or promote the interests of other persons, where one such interest is an interest in freedom. In either case, freedom is a fundamental value that provides liberal theorists with a reason for aiming to limit, control or distribute certain forms of power in certain ways. Thus, it provides a normative grounding for liberal constitutional provisions, including limited government and the separation of powers, as well as for certain economic and social policies.

The above analysis suggests the existence of a particularly strong empirical correlation between restrictions of negative freedom and those forms of social power that political liberals and republicans have traditionally been concerned to limit, control or distribute in certain ways—above all, coercion, anticipated reaction, and certain forms of conditioning and situational manipulation. Political science has it within its power to confirm or deny this correlation, and normative political theorists ought therefore to take an interest in its findings in this area.

Some republican theorists have shied away from such a reliance on falsifiable empirical correlations, preferring to establish a *logical* connection between freedom and the absence of the relevant forms of power. I do not find it helpful, however, to *define* freedom, either partly or wholly, as the absence of those forms of power with which republicans are particularly concerned (this has been the argumentative strategy adopted by Pettit (1997, 2001), and Skinner (1997, 2002). Instead, I believe it most useful to define social freedom independently of the concept of social power and then to explain why, as a matter of contingent fact, freedom (or its fair distribution) is best preserved by limiting certain forms of power or by distributing them in a certain way. For it is only on this basis that the liberal (or republican) condemnation of power as inimical to freedom will have normative force, rather than simply amounting to an analytic truth.[11] Those for whom the freedom-restricting effects of power are a logical entailment of the definition of freedom cannot cite freedom as a *reason* for wishing to control, limit or distribute power in certain ways. Defining freedom as the absence of certain forms of power wrongly assumes that the singling out of such forms of power is logically prior to an understanding of the nature of freedom. On the contrary, the logical priority should lie with our understanding of the nature of freedom.

Sometimes it is simply misguided to take refuge in the certainties of logic, when contingent empirical facts will better serve to confirm the particular structure of values we endorse. The relation between power and freedom is a case in point, and serves well to illustrate the way in which political science can help in grounding the prescriptions of political morality. In asserting a relation between certain forms of power and the unfreedom of those subject to them, liberal and republican theorists implicitly endorse a structure of values according to which an interest in limiting or redistributing power is grounded in an interest in promoting or redistributing freedom. For this reason, my own analysis of the freedom-power relation recognizes freedom as one of the fundamental, independent values in

---

[11] I present a critique of Skinner and Pettit along these lines, in part applying the analysis of the freedom-power relation contained in the present article, in Carter 2008. An earlier version of this critique can be found in Chap. 8 of Carter 1999. See also the writings of Matthew Kramer on the concept of freedom, in particular Chap. 1 of Kramer 2003, 2008.

terms of which we desire to evaluate various forms of possible social and political relation: it assumes an independently coherent conception of freedom, and then asks on this basis which of the various forms of power are accompanied by limitations of freedom. It aims to answer this last question by rendering explicit the preventive mechanisms that constantly accompany certain forms of power and not others. This constant accompaniment serves not only to explain why the liberal mind has tended, intuitively, to focus its attention on certain forms of power rather than others, but also to justify that focus in normative terms.

References

Berlin, I. (2002). Two concepts of liberty. In: I. Berlin, Liberty. Oxford: Oxford University Press.

Braham, M. (2006). Measuring specific freedom. *Economics and Philosophy, 22*, 317–333.

Braham, M., & Holler, M. J. (2009). Distributing causal responsibility in collectivities. In R. Gekker & T. Boylan (Eds.), *Economics rational choice and normative philosophy* (pp. 145–163). London: Routledge.

Carter, I. (1999). *A measure of freedom*. Qxford: Oxford University Press.

Carter, I. (2003). Positive and negative liberty. Stanford Encyclopedia of Philosophy, 2003, http://plato.stanford.edu.

Carter, I. (2008). How are power and unfreedom related? In C. Laborde & J. Maynor (Eds), *Republicanism and political theory*. NJ: Blackwell.

Carter, I., Kramer, M. H., & Steiner, H. (Eds.). (2007). *Freedom a philosophical anthology*. NJ: Blackwell.

Cohen, G. A. (1991). Capitalism, freedom and the proletariat. In D. Miller (Ed.), *Liberty*. Oxford: Oxford University Press.

Cohen, G. A. (2001). Freedom and money, Revista Argentina de Teoría Jurídica, 2, http://www.utdt.edu/departamentos/derecho/publicaciones/rtj1/pdf/finalfreedom.PDF.

Dworkin, G. (1988). *The theory and practice of autonomy*. Cambridge: Cambridge University Press.

Gabor, A., & Gabor, D. (1979). An essay on the mathematical theory of freedom. *International Journal of Social Economics, 6*, 330–371.

Gorr, M. J. (1989). *Coercion, freedom and exploitation*. Switzerland: Peter Lang.

van Hees, M. (2000). *Legal reductionism and freedom*. Berlin: Kluwer.

van Hees, M. (2003). Voluntariness, suffering, and euthanasia. *Philosophical Explorations, 6*, 50–64.

Kramer, M. H. (2003). *The quality of freedom*. Oxford: Oxford University Press.

Kramer, M. H. (2008). Liberty and domination. In C. Laborde & J. Maynor (Eds), *Republicanism and political theory*. NJ: Blackwell.

Olsaretti, S. (2004). *Liberty, desert and the market*. Cambridge: Cambridge University Press.

Oppenheim, F. E. (1961). *Dimensions of freedom. An analysis*. New York: St. Martin's Press.

Oppenheim, F. E. (1981). *Political concepts. A reconstruction*. NJ: Blackwell.

Pettit, P. (1997). *Republicanism a theory of freedom and government*. Oxford: Oxford University Press.

Pettit, P. (2001). *A theory of freedom*. Cambridge: Polity Press.

Skinner, Q. (1997). *Liberty before liberalism*. Cambridge: Cambridge University Press.

Skinner, Q. (2002). A third concept of liberty. *Proceedings of the British Academy, 117*, 237–268.

Steiner, H. (1994). *An essay on rights*. NJ: Blackwell.

Stoppino, M. (2001a). Potere e teoria politica, 3rd edn., Giuffré.

Stoppino, M. (2007). A formal classification of power. *Homo Oeconomicus, 24*, 157–181.

Stoppino, M. (2001b), What is politics? Unpublished English translation of ch. 8 of Stoppino 2001.

Taylor, C. (1979). What's wrong with negative liberty. In A. Ryan (Ed.), *The idea of freedom*. Oxford: Oxford University Press.

Taylor, M. (1982). *Community anarchy and liberty*. Cambridge: Cambridge University Press.

Causation and the Measurement of Power

Matthew Braham

Power and Cause are the same thing. Correspondent to cause and effect, are POWER and ACT; nay, those and these are the same things.

(Hobbes, *English Works*, 1, X).

1 Power Indices

The aim of this note is to elucidate the meaning of a power index in terms of causality. The usefulness of this exercise is twofold: on the one hand it casts new light on what it is that a power index measures [it tells us more than the very general notion that a power index is a direct quantification of an 'ability' called 'voting power' or as a 'reasonable expectation' of possessing this ability (Holler 1998)]; and on the other it opens up new uses for power indices, such as the measurement of responsibility in collective undertakings. It also suggests that if our analysis of power relations should capture causality then we should only focus on 'minimal winning coalitions'.

As a starting point, I assume that the concept of power has a fixed core of meaning, that of the ability of an individual to 'effect outcomes' (Braham and Holler 2005). In the context of voting power, it is effecting social outcomes, or the outcome of a formal (or informal) vote. In the literature on power indices this ability is formalized as being able to turn a decisive or winning coalition into a non-decisive or losing coalition or vice versa. From this basic definition we can obtain the Shapley-Shubik index (Shapley and Shubik 1954), the Banzhaf indices (Banzhaf 1965), the Deegan-Packel index (Deegan and Packel 1978), the Johnston index (Johnston 1978), and the Public Good Index (Holler 1982).

This chapter has been published in Homo Oeconomicus 22(4). Further developments on this chapter are published in Matthew Braham and Martin van Hees (2009). "Degrees of Causation." *Erkenntnis* 71: 323–344.

M. Braham (✉)
Institute of Philosophy, University of Bayreuth, Universitätsstr. 30,
95447 Bayreuth, Germany
e-mail: matthew.braham@uni-bayreuth.de

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_3, © Springer-Verlag Berlin Heidelberg 2013

In thumbnail form, a *power index* assigns to each player of an n-person *simple game*—a game in which each coalition that might form is either all powerful (*winning*) or completely ineffectual (*losing*)—a non-negative real number which purportedly indicates a player's ability to determine the outcome of the game. This ability is a player's power in a game given the rules of the game.

In formal terms, let $N = \{1, 2, \ldots, n\}$ be the set of players. The power set $\wp(N)$ is the set of logically possible coalitions. The simple game v is characterized by the set $W(v) \subseteq \wp(N)$ of *winning coalitions*. $W(v)$ satisfies $\emptyset \notin W(v)$; $N \in W(v)$; and if $S \in W(v)$ and $S \subseteq T$ then $T \in W(v)$. In other words, v can be represented as a pair (N, W). It should be noted that v can also be described by a characteristic function, $v : \wp(n) \to \{0, 1\}$ with $v(S) = 1$ iff $S \in W$ and 0 otherwise.

Weighted voting games are a special sub-class of simple games characterized by a non-negative real vector $(w_1, w_2, \ldots, w_n)$ where w_i represents player i's voting weight and a quota of votes, q, necessary to establish a winning coalition such that $0 < q \leq \sum_{i \in N} w_i$. A weighted voting game is represented by $[q; w_1, w_2, \ldots, w_n]$.

A power ascription in a simple game is made whenever a player i has the ability to change the outcome of a play of the game. A player i who by leaving a winning coalition $S \in W(v)$ turns it into a losing coalition $S \backslash \{i\} \notin W(v)$ has a *swing* in S and is called a *decisive member* of S.

A concise description of v can be given by a set $M(v)$, which is the set of all $S \in W(v)$ for which all members are decisive (i.e. no subset of S is in $W(v)$). A member of $M(v)$ is called a *minimal winning coalition* (MWC). Further, we denote by η_i the number of swings of player i in a game v. A player i for which $\eta_i(v) = 0$ is called a *dummy* (or *null* player) in v, i.e. it is never the case that i can turn a winning coalition into a losing coalition. (It is easy to see that i is a dummy iff it is never a member of an MWC; and i is a *dictator* if $\{i\}$ is the sole MWC).

For illustrative purposes, the *Shapley-Shubik index*, which is a special case of the Shapley value for cooperative games (Shapley 1953), measures power as the relative share of pivotal ('swing') positions of a player i in a simple game v. It is assumed that all orderings of players are equally probable. It is given by:

$$\phi_i(v) =_{def} \sum_{S \in W i \in S S \backslash \{i\} \notin W} \frac{(s-1)!(n-s)!}{n!}$$

In contrast, the *absolute Banzhaf index* for a player i in a game v measures ratio of the number of swings of i to the number of coalitions in which i is a member:

$$\beta'_i(v) =_{def} \frac{\eta_i(v)}{2^{n-1}}$$

The *normalized* or *relative* Banzhaf index divides $\eta_i(v)$ by the sum of all swings.

Holler's (1982) Public Good Index (PGI) measures the share of swings in MWCs. It is given by:

$$h_i(v) =_{def} \frac{|M_i(v)|}{\sum_{j=1}^{n} |M_j(v)|}$$

where $M_i(v)$ denotes the set of MWCs that contain i.

2 Causation and Power

The key idea that the concept of power can be elucidated in terms of causation is easy to grasp.[1] Let us begin by recapping the concept of causation. A 'cause' is a relation between events, processes, or entities in the same time series, in which one event, process, or entity C has the efficacy to produce or be part of the production, of another, the effect E, such that: (a) when C occurs, E necessarily follows (*sufficient condition*); (b) when E occurs, C must have preceded (*necessary condition*); (c) both conditions (a) and (b) prevail (*necessary and sufficient condition*); (d) when C occurs under certain conditions, E necessarily follows (contributory, but not sufficient, condition). In a very rough sense, then, C is a 'cause' or 'causal factor' of E if C is, or was, relevant or 'non-redundant' in some sense of sufficiency or necessity.[2]

In circumstances in which we are concerned with outcomes brought about by agents, i.e. an agent is a 'condition' in the sense of an event, process, or entity, we must elaborate further. If C is an agent, then to say that C is a cause of E is to postulate that C has an *action* (or sequence of actions) such that the performance of these actions under stated or implied conditions will result in E and would not result if C would not perform this action (or sequence of actions). That is, in view of her strategies, C is essential (or non-redundant) for E under the specified conditions. This corresponds precisely to the definition of a swing that is given in Sect. 1, above: if $S \in W(v)$ had formed so that the proposal passes (denoted by p), then if i has a swing in S, then the presence of i in S (or rather i's choice of the action which makes her a member of S) can be said to be a non-redundant or necessary condition of p, i.e. i is a causal factor for p because, ceteris paribus, if i was not present in S, the outcome p would not have come about, but rather not-p (rejection of the proposal).

The generic structure of a power index, that of a swing, captures, therefore, the causal role of a player. This does not, however, imply that *any* power index describes the causal factors of an outcome. This insight is the central contribution of this note. If, for example, we merely restrict our attention to swings as in the

[1] For a deeper investigation as to why we should elucidate the concept of power in terms of causation and not vice versa is discussed in Chap. 4 of Mackie (1973).

[2] It is true that this naive 'textbook' presentation of causality is crude and overrides the subtleties of what a cause or a causal factor is. But it is sufficient for my purposes here. The gamut of views can be found in the introduction to Sosa and Tooley (1993), Hitchcock (2003), and Pearl (2000).

Banzhaf indices we are confronted with the fact that for any $S \in W(v)$ the condition '$S \backslash \{i\} \notin W(v)$ but $S \cup \{i\} \in W(v)$' is not in general sufficient to provide a *full* description of the players who are causal factors for the outcome which $S \in W(v)$ can assure. Consider, for instance:

Example 1 Let $N = \{a, b, c\}$ and any two player subset can pass the proposal p. The coalitions $\{a, b, c\}$ forms.

Which players are causal conditions for p in Example 1? Given that no player has a swing—if any of the players would unilaterally choose otherwise, holding the choices of the others constant—p still pertains. As determining the decisiveness of a player in this instance clearly fails to yield a conclusion about the role of the players themselves, are we merely going to say that the coalition $\{a, b, c\}$ or its members *jointly* caused p? Is it true that none of the members of $\{a, b, c\}$ can be *singly* ascribed a causal role in bringing about p?

While an answer to these questions—which intuition suggests is negative—may not seem overly important for the analysis of *a priori* voting power per se (this is probably why it has not been discussed in the literature), it can be in other contexts. Suppose the outcome of the players' decision is a war crime for which we want to attribute responsibility and punishment. Can each player really claim innocence on the grounds that 'it was not me because I could not have done it alone nor have prevented it alone'? If we reject this defence we have to find some scheme to demonstrate that each player was causally connected to the outcome otherwise it would be but 'guilt by association'—itself as morally unacceptable as allowing people off the hook on grounds of 'collective causality'. The problem comes into stark relief with the following example:

Example 2 Let $N = \{a, b, c, d\}$ and any two player subset consisting of a, b, c can pass the proposal p. The coalition $\{a, b, c, d\}$ forms.

Player d of Example 2 is clearly a dummy (or null) player and in so being is entirely redundant for any instance of p. Thus it is false to conclude that together with the other members of $\{a, b, c, d\}$ d is jointly causal for p; p's presence in any combination of players makes no contribution whatsoever to p pertaining. Here we have an important fact: $i \in S \in W(v)$ does not imply that i is a causal condition for p. In other words, to disentangle the causal conditions for p in Examples 1 and 2 we have to determine who in $\{a, b, c\}$ and $\{a, b, c, d\}$ respectively are relevant or non-redundant for p in the sense that there exists a conjunction of players such that *each* player is necessary for p. The natural way to do this is to examine the subsets of $S \in W(v)$ that are just sufficient for p, i.e. the MWCs of S: $K \subseteq S \in M(v)$. The members of these subsets are not merely jointly causal, but each are separately causal conditions for p in the sense of being *necessary* members of these subsets. In Examples 1 and 2, these minimal sets are $\{a, b\}, \{a, c\}, \{b, c\}$, so we can say that despite the fact that in both examples no player has a swing, players a, b, and c are causal conditions for p.

The logic of this scheme is that a can be ascribed as a causal condition of p because (1) a was present in an actual or hypothetical case, (2) together with b (or *c*) *a* could bring about (or prevent) *p*, and (3) *a* was necessary for the sufficient condition that brings about *p* (although the conjunction of *a* and *b* is itself not necessary for *p* because a conjunction of *b* and *c* could also bring about *p*). Or put another way, *a* is a causal condition for *p* because of the *possibility* that *a* was necessary for *p* given that *a* voted in favour of *p* and either *b* or *c* might have decided otherwise. In other words, when examining the causal condition for social outcomes the primary unit of analysis is not the individual but coalitions.

Given that (1) $W(v)$ is uniquely determined by *N* and the set of MWCs and (2) a player has power (or causal potential) in *v* if and only if she is a member of at least one MWC (otherwise she is a dummy or null player) it should not be thought that if $i \in S \in W(v)$ and is a non-dummy in *v*, then *i* is causal condition for *p*. This is only contingently true. It is true for the case of a one-man-one-vote voting rule as in Examples 1 and 2, or in which each of the conditions (players) is equally weighted; it is not true if the game is weighted as described in Sect. 1, above. To see that a non-dummy in *v* is not a causal condition in *S,* consider:

Example 3 Let $N = \{a, b, c, d\}$ with a vector of weights $w = (35, 20, 15, 15, 15)$ in alphabetic order and a quota of 51 for *p*, i.e. $v = [51; 35, 20, 15, 15, 15]$. The coalition $\{a, b, c\}$ forms.

Although *c* is a non-dummy in *v*, *c* is redundant for $\{a, b, c\}$ and therefore in this instance cannot be said to be a causal condition for *p*. However, if the coalition $\{a, b, c, d\}$ had formed in the game of Example 3, *c* would be a causal condition because $\{a, b, c, d\}$ contains a MWC in which *c* is a member, i.e. $\{a, c, d\}$.

Those who are familiar with the philosophical literature on causation will recognise that the method of ascribing causality that I have proposed is equivalent to Mackie's (1962, 1974) conception of causation. In cases where there are complex conjunctions of conditions for some event, Mackie says that a cause is 'an *insufficient* but *necessary* part of a condition which is itself *unnecessary* but *sufficient* for the result' or from the initial letters of the italicized words, an 'inus' condition. This is not the place to elaborate, but examining the subsets $K \subseteq S \in M(v)$ that contain player *i* in order to determine if *i* is a causal condition of *p* matches precisely this idea.[3]

3 Towards a New Power Measure

To complete this sketchy framework for elucidating power and the measurement of voting power in terms of causation, I want to draw attention to two features of my argument. Firstly, there is an important difference between making causal ascriptions and measuring power. My analysis has essentially turned on hypothetical cases of actual causation: that is, in Examples 1–3 I have assumed that a particular winning

[3] In the legal theory of causation the inus condition is also known as a 'necessary element of a suffcient set' (ness). See Pearl (2000, p. 314).

coalition S formed that assured p, and then set out to disentangle who was a causal factor for p. Making such an ascription is not the same as measurement: that would be a matter of counting ascriptions. I believe that the counting of causal ascriptions is what a measure of power is all about. That is, we do not say that i is 'more causal' than j when we observe that the number of instances in which i is a necessary part of a sufficient condition for p is greater than the number of such instances for j (or in Mackie's terms, i is an 'inus' condition on more occasions than j), but rather that i is more important or more *powerful* than j. This suggests a refined interpretation of a power measure: it is the aggregation of causal ascriptions and captures the relative importance of a player in causal terms. In a slogan, power is the potential to be causal in a social world.

Secondly, what is left open is the natural form of an a priori power function that captures causality. The argument presented here intimates only that if we want to capture *all* the causal factors relevant to an instance of an outcome and that this should be part of a power measure then we must restrict our attention to MWCs for the simple reason that only MWCs guarantee that we obtain a *full* description of the causal factors for any instance of the outcome of a vote. Merely examining the instances in which i is decisive while saying something about i's contribution to the outcome does not tell us anything about the contribution of other players. What matters for ascribing causal conditions is not simply *sufficiency* but *minimal sufficiency*.[4]

The problem of constructing a power function in terms of causality is that there are two natural candidates. The first is to simply count the number of instances that a player i is a member of an MWC given the definition of the game, i.e. the quantity $|M_i(v)|$. This is none other than the non-normalized version of the PGI, also known as the Public Value (Holler and Li 1995). The second candidate is to count the number of MWCs in each $S \in W(v)$ which contain i (the set $M_i(S)$) i.e. the quantity $\sum_{S \in W(v)} |M_i(S)|$. This is a new measure altogether, which for sake of staking out my claim, I denote as the 'Causal Power Measure' (CPM). I must leave the elaboration and investigation of this new measure to another occasion.

Finally, a word or two about the advantage of elucidating power in terms of causality. It not only enriches our understanding of what a power index measures but it also provides a conceptual foundation for applying the method beyond the usual domain of constitutional issues such as equal representation in weighted voting bodies. A power index based on MWCs says, for instance, something about the responsibility that decision-makers have. It is a way of giving formal substance to the idea of the 'moral community': the nexus of individuals that can affect the lives of others by being able, either alone or in concert, to effect alternate states of

[4] It is interesting to note here that this causal perspective of power coincides with Holler's (1982) argument that when we measure power only MWCs should be considered although this does not imply that only these coalitions form. See also Marc-Wogau (1962, pp. 221–224) for reasons why we should only consider '*minimal* sufficiency' and not simply 'sufficiency' when assertions about causality.

the world. The individuals who are members of MWCs make up this community and the more of these MWCs that an individual belongs to in any given game, the more responsibility she bears.

Acknowledgments I would like to thank Manfred Holler, Hartmut Kliemt, Stefan Napel, Vincent Merlin, Mark Siebel, and Frank Steffen for comments and discussions.

References

Banzhaf, J. F. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review, 19*, 317–341.
Braham, M., & Holler, M. J. (2005). The impossibility of a preference-based power index. *Journal of Theoretical Politics, 17*, 137–158.
Deegan, J., & Packel, E. W. (1978). A new index for voting games. *International Journal of Game Theory, 7*, 113–123.
Hitchcock, C. (2003). Of humean bondage. *British Journal of Philosophy of Science, 54*, 1–25.
Holler, M. J. (1982). Forming coalitions and measuring voting power. *Political Studies, 30*, 262–271.
Holler, M. J. (1998). Two stories one power index. *Journal of Theoretical Politics, 10*, 179–190.
Holler, M. J., & Li, X. (1995). From public good index to public value an axiomatic approach and generalization. *Control and Cybernetics, 24*, 257–270.
Johnston, R. (1978). On the measurement of power: Some reactions to laver. *Environment and Planning, 10*, 569–577.
Mackie, J. L. (1962). Causes and conditions. *American Philosophical Quarterly, 2*, 245–264.
Mackie, J. L. (1973). *Truth, probability, and paradox*. Oxford: Oxford University Press.
Mackie, J. L. (1974). *The cement of the Universe*. Oxford: Oxford University Press.
Marc-Wogau, K. (1962). On historical explanation. *Theoria, 28*, 213–233.
Pearl, J. (2000). *Causality: Models, reasoning, inference*. Cambridge: Cambridge University Press.
Shapley, L. S. (1953). A value for *n*-person games. In H. Kuhn & A. W. Tucker (Eds.), *Contributions to the theory of games*. Princeton University Press: Princeton.
Shapley, L. S., & Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review, 48*, 787–792.
Sosa, E., & Tooley, M. (Eds.). (1993). *Causation*. Oxford: Oxford University Press.

Part II
Voting

Models and Reality: The Curious Case of the Absent Abstention

Dan S. Felsenthal and Moshé Machover

1 Introduction

In this chapter we address some methodological issues that arose in connection with our technical work—mainly Felsenthal and Machover (1997), but see also Felsenthal and Machover (1995, 1996, 1998, 2001) and Felsenthal et al. (1998).

Virtually all the work published so far on a priori voting power has used, explicitly or implicitly, one and the same type of mathematical structure to model decision rules of voting bodies that make yes/no decisions. We shall refer to a structure of this type as a *simple voting game* (briefly, *SVG*). Some authors consider arbitrary SVGs, whereas others confine themselves to an important subclass—that of *weighted voting games* (briefly, *WVGs*). We shall assume that the reader is familiar with the definitions of these concepts, as codified by Shapley (1962) and reproduced by other authors (for example, Straffin 1982; Felsenthal and Machover 1995). We shall also assume familiarity with the definitions of the main indices used for measuring a priori voting power: the Shapley–Shubik (S-S) index, denoted by 'ϕ' (see Shapley and Shubik 1954); and the Banzhaf (Bz) index (see Banzhaf 1965). The relative version of the Bz index (normalized so that the sum of its values for all voters in any SVG add up to 1) is denoted by 'β', and the

This article was originally published in Manfred J Holler and Guillermo Owen (eds), *Power Indices and Coalition Formation*, Boston - Dordrecht -London: Kluwer Academic Publishers, 2001, pp. 87–103. It was re-published in *Homo Oeconomicus* 19:3 (December 2002), pp. 297–310. The present version was slightly revised, and some calculation errors corrected, in July 2012.

D. S. Felsenthal (✉)
University of Haifa, Haifa, Israel
e-mail: dfelsenthal@poli.haifa.ac.il

M. Machover
King's College, London, UK
e-mail: moshe.machover@kcl.ac.uk

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*, DOI: 10.1007/978-3-642-35929-3_4, © Springer-Verlag Berlin Heidelberg 2013

so-called absolute Bz index is denoted by 'β'' (see Dubey and Shapley 1979; Straffin 1982; or Felsenthal and Machover 1995).[1]

An obvious fact about the SVG set-up is that it is strictly *binary*: it assumes that in each division[2] a voter has just two options: voting 'yes' or 'no'. On the other hand, many real-life decision rules are *ternary* in the sense that they allow abstention[3] as a *tertium quid*, which may have different effects from both a 'yes' and a 'no' vote, and so cannot be assimilated to either.[4]

This raises several interrelated questions, which we shall address in the following three sections.

In Sect. 2 we shall see how writers on voting power deal *theoretically* with the issue of abstention, and consider whether their treatment is adequate. This will lead us into a discussion of a sadly neglected distinction, first mooted by Coleman (1971), between two alternative notions underlying the formal measurement of voting power.

In Sect. 3 we shall see what adjustments, if any, scholars make when *applying* the binary model to real-life situations that are essentially ternary, and how they report the facts about such ternary rules. Here we shall conduct two brief case studies of a phenomenon familiar to philosophers of science, who refer to it as 'theory-laden observation'.

In Sect. 4 we shall consider whether an adequate ternary theoretical model can be set up; and if so, whether it yields significantly different results concerning the measurement of voting power.

2 Theoretical Discussion of Abstention

Theoretical discussion of abstention is conspicuous by its almost total absence in the literature on voting power.

An exception—which, in a sense, proves this rule—is Morriss's book (2002, Chaps. 21–24), which lies outside the mainstream publications on voting power. One of the virtues of that idiosyncratic work is that it takes abstention seriously, and makes no attempt to brush it under the carpet. However, Morriss does not propose

---

1 As a matter of fact, the Bz index is essentially a re-invention of a measure of voting power that had been proposed by Penrose (1946). In the present notation, the measure proposed by Penrose was $\beta'/2$.

2 We borrow the term 'division' from English parliamentary parlance, to denote the *collective* act of a decision-making body, whereby each individual member casts a vote. Somewhat surprisingly, writers on voting power have not made this necessary terminological distinction between the collective and the individual acts, and refer to both of them as 'voting'.

3 Unless the contrary is indicated, we use the term 'abstention' in its wide sense, including an explicit declaration 'I abstain', non-participation in the division, or absence.

4 Note that even under a ternary rule, the outcome of a division is still a dichotomy: the bill in question must either be adopted or rejected, *tertium non datur*.

an index for measuring a priori voting power (which he calls 'ability') in the presence of abstentions, nor does he define a ternary analogue of the binary SVG structure. He does outline a method of measuring a posteriori voting power (which he calls 'ableness'). We shall not enter here into an assessment of the adequacy of his outline, because in this chapter we are concerned with a priori power.

As far as the mainstream literature is concerned, the only positive treatment of abstention we have come across is in Fishburn's book (1973, pp. 53–55). Fishburn considers what might be called *self-dual ternary weighted voting games*: voters are assigned non-negative weights and are asked to express a preference or indifference between two outcomes, x and y (which may, in particular, be answers 'no' and 'yes' to a given question; in this case indifference amounts to abstention). The winning outcome is that whose supporters have greater total weight than the supporters of the alternative outcome. For such decision rules Fishburn defines a straightforward generalization of the Bz index. These ternary voting games are a very restricted class; for example, it is easy to see that the decision rule of the UN Security Council (UNSC) cannot possibly be cast in this mould.[5] Also, he makes no attempt to generalize the S-S index to his ternary weighted games. Despite its limitations, Fishburn (1973) brief treatment is a very significant positive step. But as far as we know it was not developed further in the literature published in the following 20 years.[6]

How do other writers on voting power justify their practice of confining themselves to binary theoretical models, which do not admit abstention?

Banzhaf dismisses the issue of abstention with a brief remark encaved in a footnote, (cf. Banzhaf 1965, fn. 34):

> This analysis has also assumed that all legislators are voting because this is the most effective way for each legislator to exercise his power. Naturally, some may choose to exercise their power in a less effective manner by abstaining or by being absent from the legislative chamber.

Banzhaf's argument for disregarding abstentions seems to us inadequate, as we shall explain below. But at least he does not ignore the whole issue, as do other writers.

In the published literature, as far as it is known to us, we have not found any other attempt to provide theoretical justifications for disregarding abstentions. But when presenting an earlier version of this chapter at an inter-disciplinary seminar, some of the game theorists in the audience reacted rather heatedly with a somewhat more elaborate form of Banzhaf's argument, which may be paraphrased as follows.

---

[5] Note, however, that in one respect Fishburn's model is *too* general from the viewpoint of yes/no decision-making: the model allows unresolved ties, which occur if the supporters of outcomes x and y have equal total weights. But in voting on a resolution or a bill, the outcome must be either 'yes' or 'no'; see footnote 4. The normal practice is that in case of a tie the status quo remains, so that a tie is resolved as a 'no'.

[6] See however Addendum at the end of Sect. 3 below.

> The study of voting power belongs to game theory; more specifically, it is a branch of the theory of *n*-person games. Game theory is a theory of rational behaviour. Abstaining voters are not behaving rationally, because they are not using their powers to the full. Therefore such behaviour ought to be disregarded by the theory.

This argument presupposes a particular view of voting behaviour. According to this view, advocated by Shapley, the 'worth' assigned to a coalition by the characteristic function of an SVG—1 to a winning coalition and 0 to a losing coalition—is not just a formal label. Rather, according to this view the worth of a coalition *S* represents the total payoff that the members of *S* earn when *S* is the set of 'yes' voters in a division.

What is this total payoff? Shapley's answer is quite explicit: 'the acquisition of power is the payoff' (see Abstract in Shapley 1962, p. 59).

The idea is that in winning a division, the winning coalition captures a fixed purse—the prize of power—which it then proceeds to divide among its members. The formation of the winning coalition as well as the distribution of the spoils among its members are consequent upon a process of bargaining. The motivation of voting behaviour that this view assumes has been called 'office seeking' by political scientists. It is this view of voting behaviour that underlies the S-S index and provides the justification for regarding it as a measure of voting power: the voting power of a voter is conceptualized as his or her expected or estimated share in the loot of power.

Now, it is true that from this office-seeking perspective on voting, abstention may be regarded as irrational: if by voting 'yes' you can get a share of the spoils as a member of a winning coalition that acquires power, then vote 'yes'; otherwise vote 'no'. You'll never get a prize for sitting on the fence.

But there is an alternative possible motivation of voting behaviour, which political scientists have called 'policy seeking'.[7] In his perspicacious critique of the S-S index, Coleman (1971, p. 272) points out that the latter motivation is the more usual,

> ... for the usual problem is not one in which there is a division of the spoils among the winners, but rather the problem of controlling the action of the collectivity. The action is ordinarily one that carries its own consequences or distribution of utilities, and these cannot be varied at will, i.e. cannot be split up among those who constitute the winning coalition. Instead, the typical question is ... the passage of a bill, a resolution, or a measure committing the collectivity to an action.

Indeed, this seems a realistic account of voting in, say, the UNSC. Incidentally, the UNSC is one of the two examples given by Shapley (1962, p. 59) of a 'body in which the acquisition of power is the payoff'; but it is not at all clear how passing a resolution in this body amounts to acquisition of power by those voting 'yes'.

---

[7] For a discussion of these alternative motivations in a political context, see Laver and Schofield (1990, esp. Chap. 3).

Of course, as Coleman admits, there are some circumstances where voting behaviour is motivated wholly or partly by office seeking. But the more usual cases are those where policy seeking is the predominant motivation. In these cases the outcome of a division of a decision-making body—the passage of a bill or its defeat—creates a *public good*, to which each voter may attach a utility value. This value (rather than the voter's share in some fixed purse) is the payoff that the voter ought to maximize.

It is this view of voting that underlies the absolute Bz index and its variants, proposed by Penrose and Coleman, and justifies its use as a measure of voting power: here a voter's power is conceptualized as the degree to which (or the probability that) he or she is able to affect the outcome of a division.[8]

Now, from a policy-seeking perspective on voting, the argument for disregarding abstention in theory loses most of its force. There may be several reasons why a voter would wish to abstain on a given bill. One reason can be the wish to use abstention as a way of making a public statement. The voter expects to derive some benefit not only from abstaining, but *being seen to abstain*. Such abstention should perhaps be disregarded by the theory of voting power, because it depends on the propaganda advantage of abstention itself, as a kind of side payment. Note that abstention for propaganda cannot operate if voting is secret. But in our view there are also other reasons for abstaining, which operate even when voting is secret. A voter may be indifferent to the bill, because his or her interests are not affected by it in any way. (This is a reasonable motive for abstention by absence, particularly if participation in the division involves some cost). Or the arguments for and against the bill—the estimates of the payoff to the given voter in case the bill is adopted or rejected—may be so finely balanced that the voter is unable to decide one way or the other. Is it so irrational to abstain for these reasons? It is a bizarre kind of rationality that would require you to cast a 'yes' or 'no' vote even when you couldn't care less, or when you were not sure whether passage of the bill would serve your interests better than its defeat!

The study of voting power is a branch of social-choice theory. In other branches of the theory—for example, in the study of social choice functions—it is quite normal to admit individual preference rankings that are not totally ordered but rank one or more outcomes (or candidates) as coequal. It is not that questions of individual rationality are ignored: for example, it is often argued and widely accepted that non-transitive individual preference rankings ought to be disallowed, precisely on the ground that they are not rational. But to the best of our knowledge there are not many social-choice theorists who would condemn as irrational an individual voter who does not wish or is unable to choose between Tweedledum

[8] For an elaboration of Coleman's distinction between the two conceptions of voting behaviour as underpinning two kinds of power indices, see Felsenthal et al. (1998) and Felsenthal and Machover (1998, Comment 2.2.2, Sect. 3.1 and 6.1).

and Tweedledee, or even between the Walrus and the Carpenter.[9] Why should the theory of voting power be different in this respect?

3 Theory-Laden Observation

In Sect. 2 we argued that abstention can be rational and that it should be allowed as a legitimate option in the theory of voting power. This, of course, is a matter of opinion, on which some readers may well disagree with us. We now turn to matters of fact, on which presumably there ought to be no controversy: the decision rules actually operated by certain real-life voting bodies, and the way these rules are reported in the literature on a priori voting power.

In this literature, four real-life cases are used as stock examples to illustrate the application of the theory to the real world. These are the US legislature (consisting of the two Houses of Congress and the veto-wielding President), the UN Security Council (UNSC), the mechanism (enacted in 1982) for amending the Canadian constitution, and the so-called qualified majority rule applied by the European Union's Council of Ministers to matters of a certain type. The last two examples need not concern us here, as they do indeed exclude abstention as a distinct option and can therefore be treated as SVGs, at least as far as abstention is concerned.[10] Matters are quite different in the US legislature and the UNSC. The rules of these bodies do in fact treat abstention or absence as a *tertium quid*.

First, let us consider the US legislature. Article 1, Section 5(1) of the US Constitution stipulates that business in each of the two Houses of Congress can only take place if a (simple) majority of its members are present. Beyond this, the

---

[9] Social choice theorists explicitly recognize that voters may be unable to choose between two or more alternatives, and may prefer to abstain rather than select arbitrarily one of the alternatives among which they are indifferent. Thus, for example, Brams and Fishman (1983, pp. 3–4) state that one of the advantages of *approval voting*—in which each voter can cast one vote for each candidate of whom he or she approves, and the candidate who obtains the largest number of votes is elected—is that voters 'who have no strong preference for one candidate, ...can express this fact by voting for all candidates they find acceptable ...[and thus voters] who cannot decide which of several candidates best reflects their views, would not be on the horns of a dilemma. By not being forced to make a single—perhaps arbitrary—choice, they would feel that the election system allows them to be more honest. We believe this would make voting more meaningful and encourage greater participation in elections.'

[10] Amendment of the Canadian constitution requires an assent of at least two-thirds of the provinces, inhabited by at least one-half of the total population; so here abstention counts as a 'no'.

In matters that require assent of a qualified majority of the European Union's Council of Ministers, as defined by Article 148(2) of the Single European Act, abstention counts as a 'no' vote because a fixed number of votes must be obtained in order for a resolution to be carried. According to Article 148(3), in matters that require unanimity, deliberate absence (boycott) of a member is interpreted as a 'no' vote, which amounts to a veto; but abstention of a member (whether present or represented by another member) counts as a 'yes'.

Constitution leaves it to the two Houses to fix their own rules of decision on most matters. The practice is that in each House an ordinary bill (as distinct from a decision to override a presidential veto) is deemed to pass if the necessary quorum is present and a simple majority of the members *participating in the division* vote 'yes'.[11] (The Vice President, in his role as President of the Senate, has only a casting vote, which he can use to break ties.)

The US Constitution explicitly refers to members *present* in only two instances, both concerning the Senate. Thus Article 1, Section 3(6) stipulates that in cases of impeachment the Senate's decision to convict requires the assent of at least two-thirds of the members *present*. So a President could, in theory, be convicted by the assent of just over one-third of all members, against the 'no' of just under one-sixth, with just under one-half of the members absent. Similarly, Article 2, Section 2(2), stipulates that the President shall have power, by and with the advice and consent of the Senate, to make treaties, provided two-thirds of the Senators *present* concur.

In case of a presidential veto, Article 1, Section 7(2) of the Constitution stipulates that overriding the veto requires the approval of 'two-thirds of [each] House'; but it fails to specify explicitly whether this means two-thirds of *all* members or just of those participating in the division. However, the latter interpretation was upheld by the US Supreme Court on January 7, 1919 (*Missouri Pacific Railway Co. v. State of Kansas*, 248 U.S. 276). Specifically, the Supreme Court ruled:

> "House", within Article 1, Section 7, Clause 2, of the Constitution, requiring a two-thirds vote of each house to pass a bill over a veto, means not the entire membership, but the quorum by [Article 1] Section 5 given legislative power.[12]

In their opinion the justices quoted Mr Reed, Speaker of the House of Representatives, who had ruled in 1898 that:

> The question is one that has been so often decided that it seems hardly necessary to dwell upon it. The provision of the Constitution says, "two-thirds of both Houses", what constitutes a house? ...[T]he practice is uniform that ... if a quorum is present the House is constituted, and two-thirds of those voting are sufficient in order to accomplish the object.[13]

How then do writers on voting power report these well-established facts, upon which it seemed in 1898 'hardly necessary to dwell'? The astonishing answer is that they mis-represent them. As a typical example, let us quote from Alan Taylor's recent book (Taylor 1995, p. 46).

---

[11] At present there are only two types of resolution that require approval by a prescribed proportion of an *entire* house. Senate Rule XXII (as amended by Senate Resolution 4 in 1975) requires that in order to invoke cloture (and thus limit debate) at least three-fifths of *all* Senate members (that is, currently at least 60 senators) must approve. Similarly, House Rule XXVII provides that any bill before a committee longer than 30 days may be brought before the House without committee approval, if a majority of the *entire* House (that is, currently at least 218 members) sign a petition that demands such action. This rule prevents a committee or a committee chairman from 'bottling up' by failure to report a bill upon which the House desires to vote.

[12] See Supreme Court Reporter (1920, p. 93).

[13] Ibid., p. 95.

The United States Federal System There are 537 voters in this yes–no voting system: 435 members of the House of Representatives, 100 members of the Senate, the vice president, and the president. The vice president plays the role of tie-breaker in the Senate, and the president has veto power that can be overridden by a two-thirds vote of both the House and the Senate. Thus, for a bill to pass it must be supported by either:

1. 218 or more representatives and 51 or more senators (with or without the vice president) and the president.
2. 218 or more representatives and 50 senators and the vice president and the president.
3. 290 or more representatives and 67 or more senators (with or without either the vice president and the president).

This description is of course incorrect, as it disregards abstentions. Now, Taylor is by no means a particularly careless reporter—quite the contrary.[14] And he is certainly in illustrious company. Thus Shapley (1962, p. 59) states bluntly: 'For example, the 1962 House of Representatives (when voting on ordinary legislation) $= M_{437}$.' In Shapley's notation M_n is the SVG with n voters in which the winning coalitions are those having more than $n/2$ members. On the following page Shapley displays the formula

$$\text{``Congress''} = M_{101} \times M_{437},$$

which he interprets in plain words as 'majority in both houses needed to win'.[15]

An intelligent Extra-Terrestrial visitor, presented with Shapley's report on the decision rule in the US Congress (and with no other evidence) would have to conclude that in order for ordinary legislation to pass in each of the two Houses, it needs the 'yes' of over half the *membership* of each House. This is patently false.

The hapless ET would not be disabused if he, she or it read also other scholars' writings on voting power—for example, Shapley and Shubik (1954, p. 789); Brams (1975, p. 192); Lucas (1982, p. 212); Lambert (1988, p. 235); Brams, Affuso and Kilgour (1989, p. 62) and several others. All have misrepresented the decision rule of the US legislature by implying—using plain words (like Taylor) or words and symbols (like Shapley)—that a Representative or Senator who does not vote 'yes' counts as voting 'no'.

The mis-representation of the US legislature as an SVG by Shapley and Shubik (1954) is particularly tantalizing. For, in discussing the Vice President's tie-breaking function (p. 788) they are perfectly aware that an absence of a member of the Senate during a division counts as neither 'yes' nor 'no'; and they expressly state that 'in the passage of ordinary legislation, ... perfect attendance [in the Senate] is unlikely even for important issues ...'. Yet in the very next paragraph

[14] As we shall see, he is one of the two exceptional authors we were able to find who do not misrepresent the facts about the UNSC.

[15] Since 1912, the number of members of the House of Representatives has been kept fixed at 435; so the '437' in Shapley's text is attributable to the at-large representatives given in the 86th and 87th Congresses (1959–1962) to Alaska and Hawaii (which joined the US on 3 January 1959 and on 2 August 1959, respectively). Following redistricting in 1962 the number of members in the House of Representatives has been reinstated to 435 as of the 88th Congress (1963).

(p. 789), when applying their index to the US legislature, they revert to the mis-statement of the decision rule:

> It takes majorities of Senate and House, with the President, or two-thirds majorities of Senate and House without the President, to enact a bill. We take *all* [our emphasis] the members of the three bodies and consider them voting ... .

The case of the UNSC is broadly similar, but here the tale has an interesting additional twist. During the period 1945–1965 the UNSC consisted of 11 members—five permanent members and six others. In 1966 the number of non-permanent members was increased from six to 10. The (original) Article 27 of the UN Charter stated:

(1) Each member of the Security Council shall have one vote.
(2) Decisions of the Security Council on procedural matters shall be made by an affirmative vote of seven members.
(3) Decisions of the Security Council on all other matters shall be made by an affirmative vote of seven members including the concurring votes of the permanent members; ... .

In 1966, when the UNSC was enlarged, the word ‘seven’ in clauses (2) and (3) was replaced by ‘nine’. Ostensibly, the wording of Article 27(3) of the Charter implies that in non-procedural matters an explicit ‘yes’ vote by all permanent members is needed to pass a resolution. However, in practice, as of 1946 an explicit declaration ‘I abstain’ by a permanent member is not interpreted as a veto; and as of 1947 and 1950 the same applies to non-participation in the vote and absence, respectively, of a permanent member.[16] So on non-procedural matters a resolution is carried in the UNSC if it is supported by at least nine (or, before 1966, seven) members and not explicitly opposed by any permanent member. Abstention by a non-permanent member has the same effect as a ‘no’ vote; but abstention by a permanent member is definitely a *tertium quid*. The rule is therefore essentially *ternary*, and cannot be faithfully represented as an SVG. However, this impossibility does not seem to deter most of the scholars writing on voting power. As a typical mis-statement of the facts let us quote Lambert (1988, p. 230):

> The present United Nations Security Council has 15 members. There are five major powers who are permanent members plus 10 other countries whose membership rotates. Nine votes are needed for approval of an issue, and each of the five major powers has a veto. Thus passage of an issue requires the assent of the major five and four others.

Lambert then proceeds to represent the UNSC as a WVG in which each big power has weight 7, each non-permanent member has weight 1, and the quota needed for passing a resolution is 39. Not a word about abstention. Again, Lambert is in illustrious company. Shapley (1962, p. 65), writing before the enlargement of the UNSC, says:

[16] For details on the interpretation in practice of Article 27(3) of the UN Charter with respect to abstention, non-participation or absence of a permanent member, see Simma (1982, pp. 447–454) and references cited therein.

> A somewhat more surprising example, since the voting strengths are not explicit in the rules, is the United Nations Security Council. The reader will readily verify that the following weights and quota accurately [*sic!*] define the voting system, complete with vetoes:

$$B_5 \times M_{6,2} = [27; 5, 5, 5, 5, 5, 1, 1, 1, 1, 1, 1].$$

Similar mis-statements are made by Rapoport (1970, pp. 218–219), Coleman (1971, pp. 274, 283), Brams (1975, pp. 182–191), Lucas (1982, p. 196), Riker (1982, p. 52), Brams, Affuso and Kilgour (1989, p. 58) and others.[17] It almost seems as though the Social Choice fraternity lives in an ivory tower where they can read the UN Charter but not the daily press.[18]

Note that here we are no longer concerned with opinions regarding the rationality of abstention or the desirability of taking it seriously in the theory of voting power. Nor are we concerned with how the US legislature and the UNSC *ought* to make their decisions in a perfectly rational world. We are concerned with reports about the way these bodies actually do make their decisions.

How can one explain what appear to be blatant factual errors made by a whole group of eminent scholars? Astonishing as this may be, phenomena of this sort are by no means exceptional in science, according to some philosophers of science, who refer to them as 'theory-laden observation': scientists often 'see' what their theory conditions them to expect.[19] In this they are indeed like ordinary folk; theory-laden observation has been compared to the commonplace phenomenon of optical illusion: we are 'deceived' by our senses into perceiving what our experience and (usually unconscious) suppositions lead us to expect.[20]

Notice that, according to the hypothesis we are proposing here, the neglect of abstention is not attributed to stupidity or ignorance. Indeed, several of the authors mentioned above have published papers and books on various topics in the field of

---

[17] As far as the UNSC is concerned, Bolger (1993, p. 319) was probably the first to comment explicitly on the widespread mis-reporting of the decision rule; see Addendum below.

We also wish to note that Taylor (1995, p. 46), although he presents the UNSC decision rule as an example of an SVG, is nevertheless aware that this presentation is inaccurate, and adds in parentheses: 'For simplicity, we ignore abstentions.' A similar attitude is perhaps implicit in the cautious formulation by Straffin (1982, p. 269).

[18] Anyone following press or TV reports on UNSC proceedings is in a position to notice that resolutions (on non-procedural matters) are often adopted without the assent of at least one permanent member. In the period 1946–97, this happened in the case of 300 resolutions—well over 28 % of the total 1068 resolutions adopted by the UNSC. On 15 December 1973, Resolution 344 was carried by the votes of the non-permanent members, with *all five* permanent members abstaining or not participating.

In particular, the US has long made it a firm rule never to vote for any resolution condemning Israel; but occasionally such resolutions are adopted, with the US abstaining.

And some of the authors cited must be old enough to remember that the Soviet representative was absent when the UNSC decided on 7 July 1950 to involve the UN in the Korean war.

[19] Cf. Hanson (1958) and Kuhn (1962). Perhaps a more fitting term is 'theory-*biased* observation'.

[20] Cf. Gillies (1993).

social choice, in which they do recognize and discuss abstention as a distinct option. It is only in the context of the theory of voting power that they ignore abstentions or apparently forget all about them. In our view, the best explanation of this is that the binary theoretical SVG model with which they approach the facts predisposes them to become easy victims, *in this particular context*, of the mental counterpart of optical illusion.

Speaking for ourselves, we are not claiming to be cleverer, or better informed, than all those authors—among whom are some of the greatest scholars in the field. We can attest that so long as we worked within the SVG paradigm these factual misrepresentations, which we encountered in the literature, did not evoke in us more than a vague feeling of malaise. It is only *after* we had invented, partly by chance, the alternative ternary theoretical model, which does admit abstentions (see Felsenthal and Machover 1997), that we became acutely aware of that widespread distortion. Now, being equipped with this model, we suddenly realized that many of the factual reports on decision rules that one encounters in the literature on voting power are seriously flawed.

3.1 Addendum

Quite a long time after submitting our chapter for inclusion in this volume,[21] we came across Bolger's paper (1993), which we had previously overlooked. In his paper Bolger defines (N, r) games, a generalization of cooperative games, in which each player is allowed to choose one of r alternatives. For $r = 2$ the alternatives can be 'yes' and 'no', so that an $(N, 2)$ game can be an SVG. For $r = 3$, a third alternative can be abstention. Bolger then proceeds to define a generalization of the Shapley value for (N, r) games. On the very first page of his paper, as a first example of an (N, r) game, he presents the decision rule of the UNSC, which he states correctly. He then adds, in a parenthetical remark,

> It should be noted that the U.N. Security council game is often erroneously modeled as a 2-alternative, namely 'yes' or 'no', game in which an issue passes if and only if it receives 'yes' votes from all five permanent members and at least 4 nonpermanent members.

It seems to us that this lends some support to the hypothesis proposed in this section. Bolger, who has a theoretical framework that allows for abstention as a distinct option, is not only able to observe the decision rule of the UNSC without distortion, but also notices that many others had got it wrong.

On the other hand, some doubt now seems to be cast on this hypothesis by our findings in Felsenthal and Machover (2001), in which we examine accounts of the US Congress and UNSC decision rules given in introductory textbooks on

[21] This addendum was composed after this article was accepted for publication in the 2001 volume edited by Holler and Owen mentioned in the (un-numbered) footnote appearing at the bottom of the first page of this chapter, but before it was published.

American Government and International Relations. It transpires that mistaken or misleading accounts are also quite widespread in this literature, to which the hypothesis of theory-laden observation cannot apply.

4 Ternary Voting Games

In Felsenthal and Machover (1997) we define a type of structure called a *ternary voting game* (briefly, *TVG*), which is the direct ternary analogue of an SVG: in addition to the two options of voting 'yes' or 'no', each voter may exercise a third option, abstention. We assume (a priori) that voters act independently of one another, each voting 'yes', 'no' and abstaining with equal probability of $1/3$. We define appropriate generalizations or analogues of the Bz and S-S voting-power indices for TVGs and investigate some of their properties. In particular, we determine for each n the most 'responsive' TVGs (that is, those with a maximal sum of absolute Bz values) with n voters. Here we shall confine ourselves to some general remarks.

Finding the correct ternary analogue of an SVG is not difficult. Also, the definition of an absolute (and hence also relative) Bz index for such structures is quite straightforward; in this respect Fishburn's work (1973, pp. 53–55) pointed the way.

Using a more appropriate model can have a very significant effect on the numerical results. For example, using the unsuitable SVG model for the UNSC, Straffin (1982, pp. 314–315) finds that $\beta = 0.1669$ for each of the five permanent members and $\beta = 0.0165$ for each of the 10 non-permanent members. But if one calculates the relative Bz indices while viewing the UNSC, more appropriately, as a TVG, one obtains $\beta = 0.1009$ for each of the five permanent members and $\beta = 0.0495$ for each of the 10 non-permanent members. Thus the more realistic TVG model ascribes to each non-permanent member of the UNSC a much greater relative a priori voting power than does the SVG model.

It could be argued that since abstention by a non-permanent member counts in practice as a 'no' vote, these members have in effect two voting options—'yes' and 'no'; whereas only the permanent members have three distinct options. The results obtained for the UNSC according to this 'mixed' SVG/TVG model are $\beta = 0.1038$ for each of the five permanent members, and $\beta = 0.0481$ for each of the 10 non-permanent members. These results are much closer to those of the pure TVG model than to those of the pure SVG model.

Using the same mixed SVG/TVG model, we also get quite different results from those obtained by Coleman (1971) regarding the *power of the UNSC to act*. According to Coleman's definition, the *power to act* is the a priori probability that a bill will be passed. Using the (inappropriate) SVG model, Coleman finds that the power of the UNSC to act was 0.0278 in the pre-1966 period, and that it *decreased* to 0.0259 post-1966 (cf. Coleman 1971, Table 1, p. 284). In the mixed SVG/TVG model, we obtain 0.1606 for the pre-1966 period, and an *increase* to 0.5899 thereafter.

Finding the right generalization of the S-S index is less easy. The most common representation of this index for SVGs imagines all voters lining up, in a random order, to vote 'yes' until a 'pivotal' voter tips the balance and the bill in question is adopted. The value ϕ_a of the S-S index for voter a is then the probability that a is that pivotal voter. This does not provide a clue as to how the S-S index may be generalized to the ternary case. However, there is another representation—stated (without proof) by Mann (1964, p. 153)[22]—that lends itself easily and naturally to generalization. For another approach, see Bolger (1993).

While the generalization of the S-S index to TVGs is of obvious technical interest, it may be argued that it is of limited applicability. This is because, as pointed out in Sect. 2, the underlying justification of the S-S index is as a measure of a voter's expected relative share in a fixed purse, the prize of power. But in cases where voting can be regarded in this way (as office-seeking behaviour) the argument that abstention is not rational does carry some weight. In our view this issue and, more generally, the status of the S-S index for both SVGs and TVGs requires some further study.[23]

The study of voting power in situations where abstention is a distinct option is in its infancy. We believe that its further development is both interesting and useful.

References

Banzhaf, J. F. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review*, *19*, 317–343.

Bolger, E. M. (1993). A value for games with *n* players and *r* alternatives. *International Journal of Game Theory*, *22*, 319–334.

Brams, S. J. (1975). *Game theory and politics*. New York: Free Press.

Brams, S. J., Affuso, P. J., & Kilgour, D. M. (1989). Presidential power: A game-theoretic analysis. In P. Brace, C. B. Harrington, & G. King (Eds.), *The presidency in American politics* (pp. 55–72). New York: New York University Press.

Brams, S. J., & Fishburn, P. C. (1983). *Approval voting*. Boston: Birkhäuser.

Coleman, J. S. (1971). Control of collectivities and the power of a collectivity to act. In B. Lieberman (Ed.), *Social choice* (pp. 269–300). New York: Gordon and Breach Science Publishers.

Dubey, P., & Shapley, L. S. (1979). Mathematical properties of the Banzhaf power index. *Mathematics of Operations Research*, *4*, 99–131.

[22] For a proof, see Felsenthal and Machover (1996).

[23] In this connection see Felsenthal et al. (1998). It should be pointed out here that it is simply a misconception to suppose, as some authors seem to do, that the justification for using the S-S index depends on its representation in terms of permutations, so that use of this index is legitimate in cases where the order in which voters cast their votes is important. The representation in terms of permutations is just that—a representation. The S-S index depends for whatever justification it may have on its being a special case of the Shapley value, whose justification, in turn, derives not from this or that representation but from its general mathematical properties and in particular from its characterization by Shapley's axioms (see Shapley 1953).

Felsenthal, D. S., & Machover, M. (1995). Postulates and paradoxes of relative voting power: A critical re-appraisal. *Theory and Decision*, *38*, 195–229.

Felsenthal, D. S., & Machover, M. (1996). Alternative forms of the Shapley value and the Shapley-Shubik index. *Public choice*, *87*, 315–318.

Felsenthal, D. S., & Machover, M. (1997). Ternary voting games. *International Journal of Game Theory*, *26*, 335–351.

Felsenthal, D. S., & Machover, M. (1998). *The measurement of voting power: Theory and practice, problems and paradoxes*. Cheltenham: Edward Elgar.

Felsenthal, D. S., & Machover, M. (2001). *Misreporting Rules. Homo Oeconomicus*, *17*, 371–390.

Felsenthal, D. S., Machover, M., & Zwicker, W. (1998). The bicameral postulates and indices of a priori voting power. *Theory and Decision*, *44*, 83–116.

Fishburn, P. C. (1973). *The theory of social choice*. Princeton: Princeton University Press.

Gillies, D. (1993). *Philosophy of science in the twentieth century: four central themes*. Oxford: Blackwell.

Hanson, N. (1958). *Patterns of discovery*. Cambridge: Cambridge University Press.

Kuhn, T. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Lambert, J. P. (1988). Voting games, power indices, and presidential elections. *UMAP Journal*, *9*, 216–277.

Laver, M., & Schofield, N. (1990). *Multiparty government: the politics of coalition in Europe*. Oxford: Oxford University Press.

Lucas, W. F. (1982). Measuring power in weighted voting systems. In S. J. Brams, W. F. Lucas, & P. D. Straffin (Eds.), *Political and related methods* (pp. 183–238). New York: Springer.

Mann, I., & Shapley, L. S. (1964). The a priori voting strength of the electoral college. In M. Shubik (Ed.), *Game theory and related approaches to social behavior* (pp. 151–164). New York: John Wiley.

Morriss, P., (2002). *Power–A philosophical analysis* (2nd ed.). Manchester: Manchester University Press.

Penrose, L. S. (1946). The elementary statistics of majority voting. *Journal of the Royal Statistical Society*, *109*, 53–57.

Rapoport, A. (1970). *N-person game theory: Concepts and applications*. Ann Arbor: University of Michigan Press.

Riker, W. H. (1982). *Liberalism against populism: A confrontation between the theory of democracy and the theory of social choice*. San Francisco: WH Freeman.

Shapley, L. S. (1953). A value for *n*-person games. In H. W. Kuhn & A. W. Tucker (Eds.), *Contributions to the theory of games (Annals of mathematics studies 28)* (pp. 307–317). Princeton: Princeton University Press.

Shapley, L. S. (1962). Simple games: An outline of the descriptive theory. *Behavioral Science*, *7*, 59–66.

Shapley, L. S., & Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review*, *48*, 787–792.

Simma, B. (Ed.). (1982). *The charter of the United Nations–a commentary*. New York: Oxford University Press.

Straffin, P. D. (1982). Power indices in politics. In S. J. Brams, W. F. Lucas, & P. D. Straffin (Eds.), *Political and related methods* (pp. 256–321). New York: Springer.

The Supreme Court Reporter, (1920). vol. 39. St Paul: West Publishing Co.

Taylor, A. D. (1995). *Mathematics and politics: strategy, voting, power and proof.* New York: Springer.

A Test of the Marginalist Defense of the Rational Voter Hypothesis Using Quantile Regression

Serguei Kaniovski

1 Introduction

The rational voter hypothesis, initially formulated by Downs (1957) and subsequently extended by Riker and Ordeshook (1968), holds that people vote as to maximize the expected individual net benefit of voting. In the individual calculus of voting, the gains from the desired election outcome must be factored by the probability that the vote will be instrumental in bringing about this outcome. This will be the case if the vote creates or breaks an exact tie. Because the probability of this occurring is close to zero in all but the smallest of electorates, rational choice alone cannot adequately explain why so many people routinely choose to cast an ineffectual but costly vote. This is known as the paradox of voting, one of the most persistent puzzles facing the public choice theorist.

The probability of a single vote deciding the outcome of an election rises with the expected closeness of the outcome and falls with the total number of votes cast (Sect. 2, Appendix). The rational voter hypothesis therefore predicts voter turnout to be higher in small-scale elections with close outcomes. Matsusaka and Palda (1993), Blais (2000, Chap. 1) and Mueller (2003, Sect. 14.2) review numerous empirical tests of this prediction, covering a wide range of countries and elections. Despite persistent differences across countries, types of elections and electoral systems, the evidence on the impact of closeness and electoral size on voter turnout is inconclusive, especially with respect to the size.

An earlier version of this chapter has been published in HOMO OECONOMICUS 28(1/2): "Essays in Honor of Hannu Nurmi, Volume II," edited by Manfred J. Holler, Andreas Nohn, and Hannu Vartiainen.

S. Kaniovski (✉)
Austrian Institute of Economic Research (WIFO), P.O. Box 91, 1103 Vienna, Austria
e-mail: serguei.kaniovski@wifo.ac.at

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_5,

Grofman (1993) and Blais (2000) argue that the empirical relevance of the rational voter hypothesis can only be salvaged by reducing its claim. Grofman's argument has become known as the marginalist defense of the rational voter hypothesis. The point is that, even if the expected closeness of the outcome and the number of voters cannot predict the level of voter turnout, they can provide an idea of whether and how it is affected by a change in these variables. Grofman's argument draws on the correct interpretation of the existing empirical evidence, which relies on estimates of turnout regressions discussed in Sect. 2. A turnout regression can only deliver the marginal effect of the explanatory variables (closeness and size) on the dependent variable (turnout). Because the rational voter theory can only explain the marginal effect of closeness and size on voter turnout, one must take a closer look at the strength of this relationship. Grofman proposes estimating a dynamic specification, in which the change in turnout is regressed on the change in closeness and size. In this chapter I propose a different approach, in which the static specification is augmented by the more sophisticated technique of quantile regression.

Quantile regression was proposed by Koenker and Bassett (1978). It has found applications in consumer theory, finance, and environmental studies, and is becoming an increasingly popular alternative to the OLS estimation of conditional mean models.[1] Quantile regression can be used to produce a series of estimates, each for a different quantile of the conditional turnout distribution (conditioned on closeness and size). If the election with turnout τ is, say, in the tenth quantile of the turnout distribution, then ninety percent of elections in the sample have turnouts higher than τ. Since lower quantiles correspond to elections with lower turnouts, we can distinguish the impact of closeness and size in elections where turnout was high from those where it was low. Differences in the sensitivity of turnout to closeness and size convey the importance of instrumental motivations in the respective electorates. By allowing to go beyond the conditional mean effect to uncover the impact of closeness and size on the shape of the conditional turnout distribution, quantile regression can deliver results stronger than can possibly be obtained using the OLS regressions in existing empirical studies.

How much predictive power can we expect from the rational choice theory? Perhaps not very much, if we accept the possibility that rational people might vote for reasons other than instrumental ones, for example, to express their preferences or because they consider voting their civil duty.[2] Once we admit the possibility that people may be gratified by the act of voting rather than the outcome, the existence of a relationship between closeness, size and turnout becomes a moot issue. The presence of several voter motivations raises the question of which

[1] For comprehensive surveys of recent developments, see Koenker and Hallock (2001) and Koenker (2005).

[2] Books by Brennan and Lomasky (1997), Brennan and Hamlin (2000) and Schuessler (2000) provide extensive accounts of expressive motivations in mass participation. The ethical voter hypothesis was initially proposed by Riker and Ordeshook (1968). A discussion of the importance of ethical motivations can be found in Blais (2000), who provides survey evidence in its favor.

conditions promote which type of behavior. It may well be the case that small local elections or referendums with a single clear issue are conducive to instrumental voting, whereas large, mass media assisted national elections provide an attractive arena for expressive and ethical voters. Referendums are particularly well-suited for testing the rational voter hypothesis because the typical issue put on a referendum is very specific. This facilitates the judgment of the expected utility associated with the outcome, the outcome itself being less prone to distortions related to political representation, log-rolling and other forms of strategic voting behavior.

The next section reviews the methodology of the turnout regression, which is based on the probability of a single vote deciding a two-way election. Section 3 emphasizes the need for a more differentiated approach to the empirical validation of the rational voter theory and founds the choice of data. Following a brief discussion of these data in Sects. 4 and 5 presents quantile regression estimates. The last section offers some concluding remarks.

2 The Turnout Regression

Downsian rational voter hypothesis holds that a rational citizen will vote provided the expected change in utility between her preferred and alternative outcome is larger than the cost of voting[3]:

$$P(i\ is\ decisive)\Delta u_i - c_i > 0. \tag{1}$$

The expected change in utility is usually referred to as the B-term. Turnout should increase with the probability of being decisive and with the difference in utility between the alternatives, while it should decrease with the cost of voting. The difference in utility and cost of voting are difficult to measure, which leaves the probability as the key explanatory variable. In Sect. 4 I argue that in the following analysis the error of omission should be smaller than in national elections typically studied in the literature on voter turnout.

A vote is decisive when it creates or breaks an exact tie. Under the binomial assumption on the distribution of the voting poll, if p is the prior probability that a vote will be cast in favor of the first alternative, then a single vote will decide the election approximately with the probability

$$P_e \approx \frac{2\exp(-2N(p-0.5)^2)}{\sqrt{2\pi N}}. \tag{2}$$

[3] See Downs (1957, Chaps. 11–14) and, for further developments, Tullock (1967, pp. 110–114) and Riker and Ordeshook (1968).

when N is even (subscripts refer to the parity of N).[4] The decisive vote is a tie-maker in the former and a tie-breaker in the latter case. Formula (2) is Stirling's approximation of the exact probability, further simplified for p close to one half (Appendix). They show that the efficacy of a vote will rise with closeness and fall with the total number of votes cast. Either probability is the highest at $p = 0.5$, or when the expected outcome is a tie and falls rapidly as p diverges from one half. The term $(p - 0.5)^2$ is an objective measure of closeness, which I will refer to as the *quadratic measure*. It is an *ex post* measure that can only be justified by assuming rational expectations on the part of voters. If voters are in fact rational, their subjective probability forecasts should be, on average, correct, so that an objective *ex post* measure of closeness would be equivalent to its *ex ante* counterpart. In empirical applications p is represented by the actual *split* of the voting poll. Note that applied literature traditionally assumes sincere voting. A vote is sincere if it truthfully reflects the voter's preferences. Studies on the effect of informational asymmetries on voting behavior in juries show that sincere voting is not rational and cannot be an equilibrium behavior in general (Feddersen and Pesendorfer 1996).

Taking the natural logarithm of Eq. (2) leads to the following turnout regression:

$$\log(Turnout) = \beta_0 + \beta_1(p - 0.5)^2 + \beta_2 \log(N) + \beta_3(p - 0.5)^2 N + \epsilon. \qquad (3)$$

The empirical literature knows several variations to the above specifications. The above equation is typically estimated less the interaction term $(p - 0.5)^2 N$. Although the quadratic measure of closeness is the only measure consistent with the probability (2), two alternative measures of closeness are frequently used in empirical literature: the *absolute value* $|p - 0.5|$ and the *entropy measure* $-p\log(p) - (1 - p)\log(1 - p)$ proposed by Kirchgassner and Schimmelpfennig (1992). Compared to the quadratic measure, the absolute value puts more moderate weight on p's that are far from one half. The entropy measure is a positive and convex function of p. It attains a unique maximum at one half, around which the function is symmetric. This measure differs from the other two in terms of the sign of its effect on turnout, which is positive. The expected signs on the coefficients are $\beta_1, \beta_2, \beta_3 < 0$ for the quadratic and the absolute value measures, but $\beta_1 > 0$ and $\beta_2, \beta_3 < 0$ for the entropy measure. One advantage of the entropy measure is that it can be generalized in such a way as to be applicable to an election with more than two alternatives. Note that, unlike the former two measures, the entropy measure is not defined for $p = 0$ or $p = 1$, i.e. when the expected outcome is unanimous.

Two further points are worth noting. First, the above specifications imply an inverted U-shape relationship between voter turnout and split about the point $p = 0.5$, in which the probability of being decisive attains its maximum. This relationship can be tested using the following slightly more general specification

[4] For an odd N, the analogous probability is obtained by replacing N with $N - 1$ (Appendix).

$$\log(\text{Turnout}) = \beta_0 + \beta_{11}p + \beta_{12}p^2 + \beta_2 \log(N) + \beta_3(p - 0.5)^2 N + \epsilon. \tag{4}$$

Here we expect $\beta_{11} = -\beta_{12}$, $\beta_{11} > 0$. Second, neither specifications can be used to forecast turnouts, as the dependent variable is not constrained to the unit interval. The common way to address this problem is to apply the logistic transformation to the dependent variable: $\log(Turnout/(1 - Turnout))$. Unfortunately, unlike the log-linear Downsian model, the resulting specification is highly nonlinear. In Sect. 5 I test all three measures of closeness, the invested U-shape relationship between turnout and split using the alternative specification (4), and a regression with a transformed dependent variable.

2.1 Quantile Regression on Turnout

When estimated by OLS, specification (3) yields the average marginal effects of closeness on the conditional mean of voter turnout. In a semi-logarithmic specification, the marginal effect will depend on the value of the explanatory variable. The strength of the relationship between closeness (size) and turnout is summarized in the magnitude of the coefficient on that variable. Quantile regression goes beyond the conditional mean effect to uncover the impact of closeness and size on the shape of the conditional turnout distribution. By comparing the estimates for different quantiles of the conditional turnout distribution, we can differentiate the strength of the impact of closeness and size in the conditionally low and high-turnout elections, thereby exploring the heterogeneity in the relationship. Quadratic regression has several other appealing properties such as robustness against outliers, and higher efficiency for a wide range of non-Gaussian error processes.

The objective function of quantile regression minimizes an asymmetrically weighted sum of absolute deviations, instead of the sum of squared residuals. This, and the fact that the partition into conditional quantiles depends on the entire sample, makes estimating quantile regression not even nearly equivalent to running OLS regressions on subsamples of data. Formally, let $Q_\tau(y_i|x_i) = x_i'\beta_\tau$ denote the τth conditional empirical quantile function, then

$$\hat{\beta}_\tau = \underset{\beta_\tau \in \mathbb{R}^k}{\arg\min} \left\{ \sum_{i \in \{i | y_i \geq x_i'\beta_\tau\}} \tau |y_i - x_i'\beta_\tau| + \sum_{i \in \{i | y_i < x_i'\beta_\tau\}} (1 - \tau)|y_i - x_i'\beta_\tau| \right\}. \tag{5}$$

An estimate is typically found by rewriting the above optimization problem as a linear programming problem and solving it using a modified simplex, or an interior point algorithm (Koenker 2005, Chap. 6).

As is also true of OLS regressions, the quality of inference in quantile regression depends on the number of observations and the number of parameters. In the case of quantile regression, it also depends on how finely we partition the conditional turnout distribution. Choosing a fine partition could mean relying on a few extreme observations when estimating regressions for the tail quantiles.

Given the moderate sample size of 232 observations, I estimate specification (3) using quantile regression for the 10, 25, 50 (median), 75, and 90 percent quantiles, and compare them with the conventional OLS counterparts. The more parsimonious variant of the former specification without the interaction term is also tested. Finally, a test of significance of the difference between the 90 and the 10 percent quantiles is performed. Under the i.i.d assumption on the distribution of the error process, the test statistic is asymptotically distributed as χ^2 (Koenker 2005, Sect. 3.3.2).

3 Closeness, Size and Turnout

A great part of the difficulty in validating the Downsian theory using regression analysis lies in the fact that closeness and size reflect phenomena larger and more complex than the efficacy of a vote. Matsusaka and Palda (1993), Kirchgassner and Schulz (2005) and others have argued that closeness indicates the intensity of the electoral competition. Closeness will thus reflect the pressure put on the voters rather than how they perceive the efficacy of their votes.[5] A positive correlation between closeness and turnout therefore does not imply instrumental voting, but rather how well voters are mobilized.

The issue of size is even more problematic. First, different theories of why people vote have generated conflicting predictions with respect to size. Second, and more importantly, the influence of size goes far beyond the probabilistic effect on the decisiveness of a single vote. The following examples should serve to illustrate some facets of this highly complex relationship. Schuessler (2000) imputes voters with both instrumental and expressive motivations. As the expressive voter derives utility from attaching herself to a collective election outcome, her expressive benefit will be proportional to the size of the collective to which she belongs. This results in a non-monotonic relationship between size and turnout, as large electorates confer potentially large expressive benefits, but strip the vote of all power. Schuessler's theory thus offers an explanation of why the presence of expressive motivations may be responsible for the lack of definitive empirical evidence with respect to size. Another example is the non-selfish voter theory by Edlin et al. (2005). If voters have social preferences and care about the well-being of other citizens, then the expected utility of voting could be approximately independent of the size of the electorate. This is because the subjective utility associated with imposing the desired election outcome on others, while being proportional to the size of the community, is balanced by the probability that the vote is decisive.

Finally, some explanations do not assign voters any specific motivations. Barry (1970) and Aldrich (1995), for example, argue that the expected benefits and costs

[5] See also Aldrich (1995, 1997) and Schachar and Nalebuff (1999).

of voting are simply too small for the calculus of voting to be a meaningful behavioral postulate. This view is often accompanied by the claim that most voters routinely misjudge and even ignore the efficacy of their vote. The survey evidence reported in Blais (2000) to an extent corroborates this view.

It seems that at least part of the difficulty in obtaining definitive empirical evidence on the rational voter hypothesis lies in the roundabout approach taken in the literature. A direct calculation using formula (2) shows that in an electorate of just 1001 voters the probability that a single vote will be decisive cannot exceed 0.0252. The numerical smallness of the direct probability measure poses a great empirical difficulty. As the discussion in the previous section indicated, a common way of circumventing this problem is to separate the probabilistic effect of closeness from that of size. One drawback of doing this is that, taken separately, closeness and size will pick up effects quite unrelated to the efficacy of the vote. The larger and more significant the election, the more distorted the relationship between closeness, size, and turnout are likely to be. Using data for voter turnouts in Norwegian school language referendums, Kaniovski and Mueller (2006) have tested an alternative explanation of why the size of the electorate may reduce turnout. Large communities are, on average, more heterogeneous. From the literature on community participation surveyed in Costa and Kahn (2003) we know that the willingness to participate decreases with heterogeneity. The detrimental effect of heterogeneity on participation in general, and turnout in particular, may compound the probabilistic effect of size on the decisiveness of the vote. It therefore comes as no surprise that the size of the electorate has little explanatory power in large elections, such as national presidential or legislative elections, or in countrywide referendums. The larger the electorate, the more distorted we believe the relationship between closeness, size, and turnout will be. This must be especially true with respect to size.

4 The Data

For closeness and size to have a reasonable explanatory power, we need to turn our attention to small electorates, in which pronounced instrumental motivations can realistically be expected. Furthermore, the majority of empirical studies derive specifications based on the probability that one vote will decide an election with only two alternatives discussed in Sect. 2. Both considerations point to local referendums as the best source of data for testing the rational voter hypothesis. The turnout record in 232 school district referendums in Norway ideally fulfills the smallness and the binary choice criteria, and has already been used in Søberg and Tangerås (2004) to test the rational voter hypothesis, as well as in Kaniovski and Mueller (2006) to study the effect of heterogeneity on voter turnout.

On 232 occasions between 1971 and 2003, Norwegians were asked which of the two official languages, Bokmål or Nynorsk, should be the primary language of their school district. With relatively small electorate sizes, ranging from 6 to 4,625

Table 1 Descriptive statistics

	Obs.	Mean	St.Dev.	Min.	Max.
Turnout	232	0.66	0.23	0.10	1.00
Electorate	232	394.98	561.48	6	4625
Split	232	0.47	0.18	0	0.89
Measures of closeness					
Quadratic measure	232	0.03	0.05	0	0.25
Absolute value	232	0.14	0.12	0	0.50
Entropy measure	229	0.63	0.11	0.05	0.69

The split is defined as the ration of votes in favor of Nynorsk to the total number of votes cast. The entropy measure is not defined for unanimous outcomes, hence fewer observations

and an average of 395 voters, these referendums fulfill the smallness criterion while still offering sufficient variability for robust empirical inference, covering more than three decades and 76 municipalities in 13 of Norway's 19 counties (Table 1).

Søberg and Tangerås (2004) estimate an OLS regression using the absolute value as the measure of closeness. They find both the size of electorate in the school district and the expected closeness of the outcome to be good predictors of voter turnout. Prior to 1985, the referendums were semi-binding (binding, provided that at least 40 percent of the electorate voted in favor). All 86 referendums since 14.06.1985 have only been advisory, although the outcomes of all except four of the advisory referendums were implemented by the municipal authorities. Participation in some referendums was limited to the parents of school children.[6] Søberg and Tangerås (2004) show that both circumstances have had their predicted effect on voter turnout, which has also been confirmed by Kaniovski and Mueller (2006). Higher turnouts in semi-binding referendums reflect the fact that a vote is more decisive in this type of referendum. Extending the franchise to parents only further increased turnout, after controlling for the size effect, presumably because the parents of school children were more concerned with the issue than the general public. These previous findings illustrates the importance of subjective utility derived from the desired election outcome in these referendums, which is the B-term in the Downsian model.

Consistent with the expected utility maximization, the decision to vote will depend on this utility, which is the B-term in the Downsian model. Although it is virtually impossible to capture the B-term empirically, we shall expect that any collective outcome will bring different subjective utilities to different people. This heterogeneity will rise with the size of the community and the number and complexity of the issues, and could well be responsible for the empirical difficulties mentioned in the introduction.

The relative simplicity of the issue at hand suggests that voters derive roughly similar subjective utilities, which can therefore be omitted from the analysis. In a

[6] This was the case from 01.07.1971 to 31.07.1985, and from 01.08.1999 to 31.07.2000.

series of national legislative elections where several parties are pushing a variety of issues—some openly, others covertly, the heterogeneity of subjective utilities is likely to be much higher than in a series of school language referendums. The above considerations make Norwegian school-district referendums an attractive choice for testing the rational voter hypothesis.

5 Quantile Regression Results

I begin by estimating quantile regressions for the three specifications—one for each measure of closeness—for the 10, 25, 50, 75, and 90 percent quantiles of the conditional turnout distribution. To control for the fact that a vote in a semi-binding referendum is more decisive than in an advisory referendum, a dummy variable discriminating between the two legal settings is included. I do not control for "parents only" ruling, as its effect on decisiveness is already reflected in the smaller size of the electorate, and I am primarily concerned with the precise measurement of the effect of closeness and size. Table 2 compares quantile regression estimates to the corresponding OLS estimates contained in the second column.

Results reported in Table 2 show that the coefficients on closeness and size have their expected signs in all regression and are mostly significantly different from zero, all variables together explaining between 20 and 60 percent of the observed variation in turnout. The factors that increase the efficacy of the vote also increase voter turnout. This would be our conclusion even if we were confined to OLS (the second column in Table 2). It is equally apparent, however, that the OLS regression obscures important detail. First, the effects of closeness and size vary for different portions (quantiles) of the conditional turnout distribution and, second, this variation has a pattern. Since lower conditional quantiles correspond to the referendums with lower turnouts, the results show that low-turnout referendums have had a weaker positive impact of closeness and a stronger negative impact of the electorate size, the opposite being true of high-turnout referendums. In other words, the disparity between turnouts in close and clear-cut referendums is substantial, particularly at the left tail of the conditional distribution, and this disparity decreases nearly monotonically as we move to the right tail of the distribution. A similar statement can also be made for the size. This further substantiates the prediction of the Downsian model.

The differences in the sensitivity of turnout to closeness and size increase nearly monotonically across quantiles. To test whether the differences in the strength of the relationship between closeness, size and turnout are statistically significant, I test for the difference in the 10 and 90 percent quantiles estimates. The last column in Table 2 reports the magnitude of these differences, which are indeed significant. Not only does this test confirm the predicted effect of closeness and size, it also shows these variables to cause the dispersion in observed turnout levels. Interestingly, the coefficients on the dummy variable do not show a

Table 2 Quantile regression

Quantile in percent[a]	OLS[b]	10	25	50	75	90	90–10
Dependent variable	Log(Turnout)						
Quadratic measure	−2.226 (−4.26) ***	−3.266 (−2.70) ***	−2.755 (−3.18) ***	−1.225 (−1.85) *	−0.508 (−0.93)	−0.821 (−1.51)	2.445 (1.93) *
Log(Electorate)	−0.206 (−6.84) ***	−0.276 (−4.58) ***	−0.202 (−3.35) ***	−0.135 (−5.40) ***	−0.100 (−3.79) ***	−0.103 (−4.42) ***	0.172 (2.77) ***
Inter. Term	−0.004 (−3.60) ***	−0.006 (−2.12) **	−0.005 (−2.25) **	−0.007 (−4.05) ***	−0.007 (−2.62) ***	−0.001 (−0.45)	0.004 (1.11)
Semi-binding = 1	0.132 (2.49) **	0.181 (1.10)	0.281 (2.20) **	0.180 (3.44) ***	0.159 (3.23) ***	0.108 (3.37) ***	−0.074 (−0.45)
R^2	0.59	0.45	0.42	0.36	0.27	0.23	
Absolute value	−0.908 (−4.13) ***	−1.668 (−4.92) ***	−1.244 (−3.13) ***	−0.387 (−1.61)	−0.065 (−0.36)	0.015 (0.09)	1.682 (4.55) ***
Log (Electorate)	−0.169 (−6.08) ***	−0.280 (−4.81) ***	−0.191 (−3.50) ***	−0.100 (−3.91) ***	−0.070 (−3.19) ***	−0.057 (−2.37) **	0.224 (3.69) ***
Inter. Term	−0.001 (−4.78) ***	−0.001 (−1.30)	−0.002 (−3.04) ***	−0.002 (−3.52) ***	−0.002 (−3.05) ***	−0.002 (−2.72) ***	−0.001 (−1.10)
Semi-binding = 1	0.125 (2.37) **	0.133 (0.95)	0.175 (1.62)	0.204 (4.20) ***	0.140 (2.76) ***	0.061 (1.89) *	−0.072 (−0.53)
R^2	0.58	0.46	0.42	0.34	0.26	0.24	
Entropy measure	1.754 (8.56) ***	2.802 (4.78) ***	2.288 (6.26) ***	1.732 (5.21) ***	0.702 (1.86) *	0.435 (2.38) **	−2.367 (−3.94) ***
Log(Electorate)	−0.174 (−4.16) ***	−0.186 (−2.24) **	−0.138 (−2.54) **	−0.063 (−1.60)	−0.092 (−2.06) **	−0.084 (−2.63) ***	0.103 (1.20) *
Inter. Term	−0.000 (−2.06) **	−0.000 (−1.29)	−0.001 (−3.91) ***	−0.001 (−3.18) ***	−0.000 (−1.24)	−0.000 (−1.10)	0.000 (0.63)
Semi-binding = 1	0.095 (1.81) *	0.100 (0.69)	0.116 (1.23)	0.146 (3.33) ***	0.107 (2.07) **	0.099 (3.03) ***	−0.001 (−0.01)
R^2	0.60	0.46	0.46	0.37	0.27	0.25	

[a]Bootstrap Standard Errors, Pseudo R^2

[b]Robust Standard Errors

*** 1 percent; ** 5 percent; * 10 percent level of significance; The estimate for the constant term is omitted

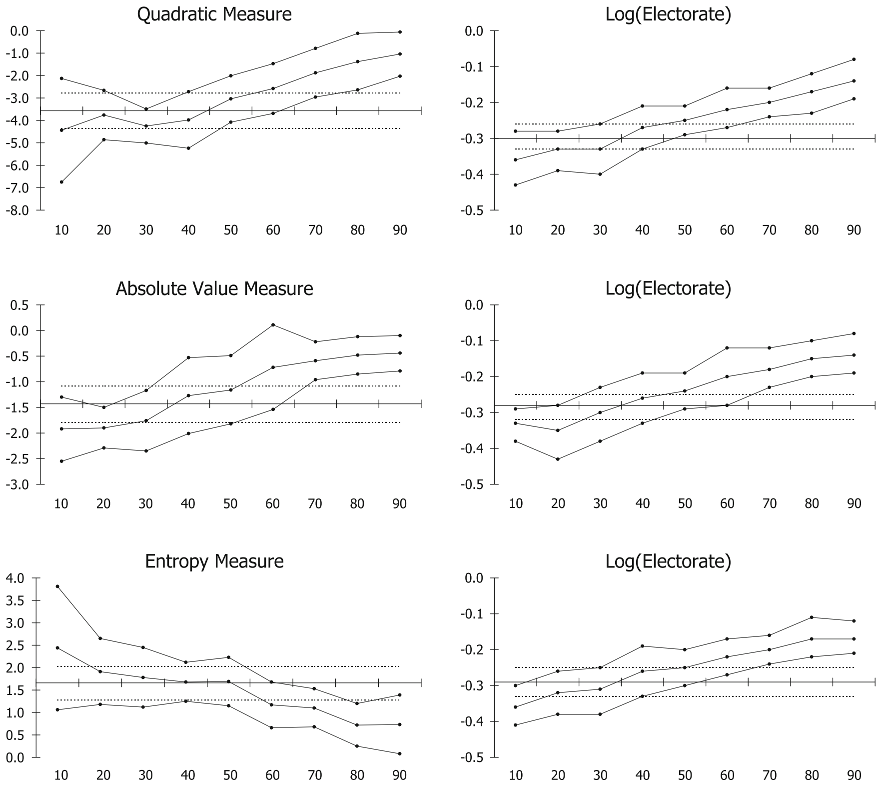


Fig. 1 Quantile regression results for the three measures of closeness. The horizontal axis crosses the vertical axis at the OLS estimate, whose 95 percent confidence interval is indicated with *hatched lines*

monotonic pattern, despite the fact that most coefficients are highly significant and have their predicted signs. The type of the referendum has had a more uniform effect on voter turnout in these referendums.

As a further test, I estimate a more parsimonious specification, one without the interaction term and the dummy variable. This time I estimate quantile regression for the $10, 20, \ldots, 90$ percent quantiles (the deciles) for the three specifications. Figure 1 plots the coefficients on closeness and size for the deciles of the conditional turnout distribution. Estimates for the consecutive quantiles and their 95 percent confidence intervals are connected by solid lines. To facilitate comparison, the horizontal axis is centered on the OLS estimate, its 95 confidence interval shown with hatched lines. All coefficients have their expected signs and are significantly different from zero in every regression. Results of the parsimonious specification corroborate those reported in Table 2, sowing a surprisingly clear nearly monotonic pattern. A lower observed turnout is again accompanied by a weaker positive effect of closeness and a stronger negative effect of size, all results being robust to the choice of closeness measure.

Table 3 Quantile regression

Quantile in percent[a]	OLS[b]	10	25	50	75	90	90–10
Dependent variable				Log(Turnout)			
Split	2.157	4.510	2.939	1.133	0.480	0.817	−3.694
	(4.05)	(3.48)	(3.26)	(1.63)	(0.82)	(1.44)	(−2.67)
	***	***	***				***
Split^2	−2.070	−5.093	−2.996	−1.016	−0.378	−0.648	4.445
	(−3.51)	(−3.21)	(−2.78)	(−1.61)	(−0.62)	(−1.10)	(2.68)
	***	***	***				***
Log(Electorate)	−0.207	−0.291	−0.209	−0.136	−0.105	−0.125	0.165
	(−6.83)	(−5.09)	(−3.51)	(−5.11)	(−4.13)	(−6.16)	(2.73)
	***	***	***	***	***	***	***
Inter. Term	−0.005	−0.004	−0.005	−0.007	−0.007	−0.002	0.002
	(−3.66)	(−1.48)	(−2.18)	(−4.09)	(−2.56)	(−0.63)	(0.64)
	***		**	***	**		
Semi-binding = 1	0.131	0.088	0.281	0.193	0.156	0.082	−0.006
	(2.48)	(0.60)	(2.27)	(3.73)	(3.37)	(3.00)	(−0.04)
	**		**	***	***	***	
R^2	0.59	0.46	0.42	0.36	0.27	0.24	
Dependent variable				Log(Turnout/(1-Turnout))			
Split	6.973	8.684	9.094	7.911	5.255	4.54	−4.145
	(4.04)	(3.43)	(4.21)	(3.01)	(1.57)	(1.34)	(−1.03)
	***	***	***	***			
Split^2	−6.739	−9.242	−9.100	−7.719	−5.005	−3.891	5.351
	(−3.69)	(−2.97)	(−3.40)	(−2.96)	(−1.49)	(−1.16)	(1.22)
	***	***	***	***			
Log(Electorate)	−0.622	−0.585	−0.608	−0.565	−0.565	−0.766	−0.181
	(−7.91)	(−4.03)	(−4.83)	(−5.31)	(−4.64)	(−8.44)	(−1.13)
	***	***	***	***	***	***	
Inter. Term	−0.003	0.000	−0.002	−0.002	−0.005	0.001	0.001
	(−1.21)	(0.01)	(−0.53)	(−0.54)	(−0.98)	(0.10)	(0.08)
Semi-binding = 1	0.412	0.300	0.489	0.511	0.597	0.560	0.260
	(3.01)	(1.10)	(2.15)	(3.68)	(3.00)	(3.40)	(0.86)
	***		**	***	***	***	
R^2	0.52	0.36	0.38	0.34	0.30	0.31	

Alternative Specifications
[a]Bootstrap Standard Errors, Pseudo R^2
[b]Robust Standard Errors
*** 1 percent; ** 5 percent; * 10 percent level of significance; The estimate for the constant term is omitted

The results from the two alternative specifications presented in Table 3 closely resemble those of the basic specification with the quadratic measure of closeness at the top of Table 2. When entered separately, *Split* and *Split*^{2} produce estimates of opposite signs and similar absolute values. With $F(1, 226) = 0.29$ for the first $F(1, 221) = 0.36$ for the second OLS specification, a Wald-test of the linear restriction $\beta_{11} + \beta_{12} = 0$ indicates no difference in the absolute values of the two coefficients. This is also true for the individual quantile estimates. For example, the test statistics for the 10 percent quantile estimates are, respectively, $F(1, 226) = 1.60$ and $F(1, 221) = 0.42$. In sum, alternative specifications indicate the robustness of the basic turnout regression and the existence of an inverted U-shape relationship between voter turnout and split, as predicted by the Downsian model.

6 Summary

The marginalist defense of the Downsian rational voter hypothesis asserts that, while closeness and size cannot explain the absolute level of turnout, they can account for change in these variables. In this chapter I show that a regression analysis more sophisticated than that hitherto employed in the literature can add further weight to the marginalist cause.

The novelty of this study lies in its use of quantile regression to investigate the heterogeneity in the strength of the relationship between closeness, size and turnout. Quantile regression reveals the impact of closeness and size on the shape of the conditional turnout distribution and thus delivers results stronger than can possibly be obtained using OLS regressions in existing empirical studies.

Survey evidence tells us that voters are driven by several distinct motivations and this urges us to consider which conditions promote which type of behavior. It seems reasonable that large, mass media assisted national elections may be an attractive arena for the expressive and the ethical voter, while small local elections or referendums with a single clear-cut issue may be more conducive to instrumental voting.

The empirical results presented in this chapter support the second hypothesis. Whatever caused the differences in turnout in the 232 Norwegian school language referendums, they can to a large extent be explained by factors relating to instrumental voting. Quantile regression shows that a lower observed turnout is accompanied by a weaker positive effect of closeness and a stronger negative effect of size, with the differences being significant and robust to the choice of closeness measure. This pattern corroborates the average marginal effect uncovered by OLS. Both findings support the marginalist defense.

Acknowledgments I would like to thank Morten Søberg, who provided the data on Norwegian school language referendums, and Dennis C. Mueller for constructive criticism.

Appendix Probability that a Vote is Decisive in a Two-Way Election

When a voter faces two alternatives her vote becomes decisive either when all other votes would have tied the outcome (Event 1), or when her preferred alternative would lose by a single vote if she abstained (Event 2). The two events are mutually exclusive, as N is odd in the former case and even in the latter. Let p be a prior probability of a vote being cast for the voter's preferred alternative. Event 1 occurs with probability P_o, which is the probability of $\frac{N-1}{2}$ successes in $N-1$ Bernoulli trials with the probability of success p:

$$P_o = \frac{(N-1)!}{(\frac{N-1}{2}!)^2} p^{\frac{N-1}{2}} (1-p)^{\frac{N-1}{2}}. \tag{6}$$

Since N is odd, substitute $N = 2k+1$ for $k = 0, 1,$

$$P_o = \frac{(2k)!}{(k!)^2} p^k (1-p)^k. \tag{7}$$

By the Stirling's approximation $x! \cong \sqrt{2\pi}(x^{x+0.5}e^{-x})$, where $\cong$ means that the ratio of the right hand side to the left hand side approaches unity as $x \to \infty$,

$$P_o \cong \frac{\sqrt{2\pi}(2k)^{2k+0.5}e^{-2k}}{(2\pi)(k^{k+0.5}e^{-k})^2} p^k (1-p)^k = \frac{2^{2k+0.5}}{\sqrt{2\pi k}} p^k (1-p)^k. \tag{8}$$

Substituting back $k = (N-1)/2$ yields, after some simplification,

$$P_o \approx \frac{2[1-(2p-1)^2]^{\frac{N-1}{2}}}{2\sqrt{\pi(N-1)}}. \tag{9}$$

Note that in $[0, 1]$ both $x(1-x)$ and $1-(2x-1)^2$ attain their maxima at $x = 0.5$, so that the approximation preserves P_o's essential property of being highest at $p = 0.5$. Using the fact that $1 + x \approx e^x$ for small $|x|$ and $1-(2p-1)^2 \approx e^{-(2p-1)^2} = e^{-4(p-0.5)^2}$, for all p close to 0.5 the above expression can written as

$$P_o \approx \frac{2e^{-2(N-1)(p-0.5)^2}}{\sqrt{2\pi(N-1)}}. \tag{10}$$

This formula leads to the convenient log-linear specification with an interaction term between the quadratic measure of closeness $(p-0.5)^2$ and size N.

Event 2 occurs with probability P_e, which is the probability of $\frac{N}{2}$ successes in $N-1$ Bernoulli trials with the probability of success p. By a similar argument using the parity of N, for all p close to 0.5,

$$P_e \approx \frac{2e^{-2N(p-0.5)^2}}{\sqrt{2\pi N}}. \tag{11}$$

Good and Mayer (1975) discuss the magnitude of error in P_o and P_e due to p deviating from 0.5, which can be substantial (Fig. 2). See, also Chamberlain and Rothschild (1981), and in the context of voting power, Grofman (1981). Kaniovski (2008) computes the probability of casting a decisive vote when votes are neither

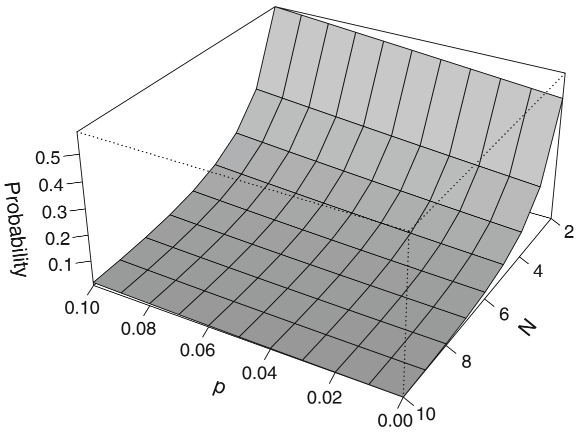


Fig. 2 Approximation to the probability of casting a decisive vote. In an election with two alternatives, the *Probability* of being decisive decreases with the size of electorate N and with $|p - 0.5|$. For a fixed N, the probability is the highest when $p = 0.5$ and decreases rapidly as p deviates from 0.5. The approximation is valid for $p \approx 0.5$

equally probable to be for or against, nor independent. Departures from either assumption induce a substantial bias in this probability compared to the baseline case of equally probable and independent votes. The bias incurred by the probability deviating from one-half is larger than that incurred by the Pearson product-moment correlation coefficient deviating from zero.

References

Aldrich, J. H. (1995). *Why parties? The origin and transformation of political parties in America*. Chicago: University of Chicago Press.

Aldrich, J. H. (1997). When is it rational to vote? In D. C. Mueller (Ed.), *Perspectives on public choice* (pp. 373–390). Cambridge: Cambridge University Press.

Barry, B. (1970). *Sociologists, economists, and democracy*. London: Collier Macmillan Publishers.

Blais, A. (2000). *To vote or not to vote: The merits and limits of rational choice theory*. Pittsburgh: University of Pittsburgh Press.

Brennan, H. G., & Hamlin, A. (2000). *Democratic devices and desires*. Cambridge: Cambridge University Press.

Brennan, H. G., & Lomasky, L. E. (1997). *Democracy and decision*. Cambridge: Cambridge University Press.

Chamberlain, G., & Rothschild, M. (1981). A note on the probability of casting a decisive vote. *Journal of Economic Theory*, *25*, 152–162.

Costa, D. L., & Kahn, M. E. (2003). Civic engagement and community heterogeneity: An economist's perspective. *Perspectives on Politics 1*, 103–111. Unpublished manuscript, Harvard University.

Downs, A. (1957). *An economic theory of democracy*. New York: Harper & Row.

Edlin, A., Gelman, A., & Kaplan, N. (2005). *Voting as a rational choice: The effect of preferences regarding the well-being of others*. Unpublished manuscript, University of California, Berkeley.

Feddersen, T., & Pesendorfer, W. (1996). The swing voters curse. *American Political Science Review*, *86*, 408–424.

Good, I., & Mayer, L. (1975). Estimating the efficacy of a vote. *Behavioral Science*, *20*, 25–33.

Grofman, B. (1981). Fair apportionment and the Banzhaf index. *The American Mathematical Monthly*, *88*, 1–5.

Grofman, B. (1993). Is turnout the paradox that ate rational choice theory? In B. Grofman (Ed.), *Perspectives on public choice* (pp. 93–103). University of Michigan Press.

Kaniovski, S. (2008). The exact bias of the Banzhaf measure of power when votes are neither equiprobable nor independent. *Social Choice and Welfare*, *31*, 281–300.

Kaniovski, S., & Mueller, D. C. (2006). Community size, heterogeneity and voter turnouts. *Public Choice*, *129*, 399–415.

Kirchgässner, G., & Schimmelpfennig, J. (1992). Closeness counts if it matters for electoral victory: Some empirical results for the United Kingdom and the Federal Republic of Germany. *Public Choice*, *73*, 283–299.

Kirchgässner, G., & Schulz, T. (2005). *Expected closeness or mobilisation: Why do voters go to the polls?—empirical results for Switzerland, 1981–1999*. Unpublished manuscript, University of St.Gallen.

Koenker, R. (2005). *Quantile regression*. New York: Econometric Society Monographs, Cambridge University Press.

Koenker, R., & Bassett, G. W. (1978). Regression quantiles. *Econometrica*, *46*, 30–50.

Koenker, R., & Hallock, K. F. (2001). Quantile regression. *Journal of Economic Perspectives*, *15*, 143–156.

Matsusaka, J. G., & Palda, F. (1993). The Downsian voter meets the ecological fallacy. *Public Choice*, *77*, 855–878.

Mueller, D. C. (2003). *Public choice III*. Cambridge: Cambridge University Press.

Riker, W. H., & Ordeshook, P. C. (1968). A theory of the calculus of voting. *American Political Science Review*, *62*, 25–42.

Schachar, R., & Nalebuff, B. (1999). Follow the leader: Theory and evidence on political participation. *American Economic Review*, *89*, 525–549.

Schuessler, A. A. (2000). *A logic of expressive choice*. Princeton: Princeton University Press.

Søberg, M., & Tangerås, T. P. (2004). *Rational voter turnout*. Evidence from 232 Norwegian referendums. Unpublished manuscript, The Research Institute of Industrial Economics, Stockholm.

Tullock, G. (1967). *Towards a mathematics of politics*. Ann Arbor: University of Michigan Press.

Intensity Comparisons, the Borda Rule and Democratic Theory

Eerik Lagerspetz

1 The Issue

This article may be seen as a small contribution to a larger project, an attempt to link the theory of social choice to more traditional normative political philosophy[1]. Although the subject of this chapter is applied social choice, I am writing as a philosopher: I try to reflect some well-established results rather than to prove new ones. More specifically, I shall discuss the consequences of the common idea that our decision-making methods should take the intensities of preference into account. In the theory of social choice, the possibility of making interpersonal intensity comparisons is often seen as the way out from Arrow's problem. In ethics, the relevancy of such comparisons has been defended in terms of utilitarianism as well as in terms of fairness. Finally, in political science it is connected to the discussion on various decision-making mechanisms. These discussions overlap, but there are very few attempts to bring them together in a systematic way.

The theory of social choice can be applied to different contexts. Here are some examples: evaluation in ethics and welfare economics, voting in democratic bodies and in elections, decision-making in courts and panels of experts, multiple-criteria decision-making in planning, engineering, and quality assessment, choosing the

An earlier version of this chapter was published in *Essays in Honor of Hannu Nurmi* (Homo Oeconomicus 28:49–70), 2011.

[1] This text is a slightly revised version of a chapter (with the same title) published in *Homo Oeconomicus* 28 (2011), pp. 49–70. I am grateful to the two anonymous referees of the journal for their valuable criticism, and to Hannu Nurmi and Manfred J. Holler for useful comments.

E. Lagerspetz (✉)
Department of Behavioural Sciences and Philosophy, University of Turku,
FI 20014 Turku, Finland
e-mail: eerlag@utu.fi

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_6,

winners in various contests of skill, aggregating information in opinion measurements and marketing research. All processes, rules and theories that are purported to select an alternative or several alternatives or to produce an overall ranking by using individual rankings as the starting-point may be interpreted in terms of the theory of social choice. In this article, I focus the context of democratic decision-making, taking the fundamental democratic values—voter equality, voters' effective influence ('popular sovereignty') and freedom of choice—as granted.[2] The critical question is, then, whether it is possible to find an institutional method which would deliver the required information about voters' preference intensities while satisfying the requirements of democracy. Numerous theoretical proposals has been made (see, for example, Hillinger 2004, 2005), but the only method which has actually been used in political contexts is the Borda rule (or Borda count). I discuss the two most sophisticated defences of that method, those presented by Michael Dummett and by Donald Saari. While the arguments put forth by Dummett and especially by Saari are theoretically convincing, I shall argue that matters tend to be more complex when Borda-like systems are actually applied in democratic decision-making. I try to show that the arguments for the Borda rule are partly dependent on the view that voting rules are means to *acquire information* about voters' preferences. Voting rules, however, do not aggregate preferences. They aggregate *votes* which are more or less truthful expressions of voters' preferences between those alternatives which happen to be on the agenda.

One of the important aspects of the formal theory of social choice is that it can be applied to different contexts. However, the fundamental problem often neglected by the theorists of social choice is that while the formal apparatus may be applied to all kinds of aggregation processes, different considerations may be relevant in different contexts. In political contexts, there are two aspects which are not equally relevant in some other cases: the requirements of democracy, and the interaction between the choice of an aggregation method (voting rule) and the input of aggregation (votes cast). Moreover, contrary to many theorists, I do not think that the interaction problem is solved by supposing that all voters are fully rational, always having complete preference rankings and acting in the strategically optimal way. Instead of seeing voting as one process of information-aggregation among others, we should, perhaps, see it primarily as an exercise of power. This power should, like all power, be constrained by normative rules. The theory of social choice is able to capture some, but only some, part of the normative aspect of voting. While the arguments for the use of Borda-like rules may be convincing for example in multi-criteria decision-making, they need not to be equally convincing in voting contexts.

[2] The discussion on the fundamental values of democracy has to be saved to another occasion. There is a vast amount of literature on the subject; for an interested reader, I may recommend the books of Christiano (1996) and Saward (1998) as good starting points. (As the reader may guess, the position taken in those works is quite close to my own.)

2 Arguments for and Against Intensity Comparisons

If we could compare different decision alternatives in terms of the intensity by which they were supported or opposed, our collective decisions would not need to be based solely on ordinal rankings. There are at least three possible reasons intensity comparisons as relevant. First, their relevancy follows from the general utilitarian programme. Second, most notions of fairness presuppose some forms of interpersonal comparisons at some level. In democratic theory, the problem of "intense minorities" is usually seen as a problem of fairness, not of maximization (for overviews of the problem, see Dahl 1956, pp. 48–50, 90–102; Kendall and Carey 1968; Jones 1988; Karvonen 2004). Third, intensity comparisons seem to provide an escape-route from Kenneth Arrow's famous impossibility result. One possible way of interpreting Arrow's Theorem is that an interest-based political theory like utilitarianism cannot be based on ordinal comparisons. In order to define the common good or general interest, we need some additional information. If we are utilitarians, we either have to reject the whole idea that decisions should be based on individual preferences, or we have to endorse full-blown utilitarianism with an interpersonally applicable measure of intensities (Ng 1979). We should be able to say, in a truly utilitarian fashion, that an alternative is so-and-so many units better than another alternative when measured on some absolute scale. The question is how to get reliable information about these differences.

Conversely, it is possible distinguish at least three reasons for *rejecting* interpersonal intensity comparisons in voting contexts. First, some theorists—following the famous critique made by the economist Lionel Robbins in 1932—regard such comparisons as *conceptually* meaningless. Even if voters were allowed to express their preferences in cardinal terms, the numbers would not measure intensities. Intensities are not observable; judgments about intensities are necessarily based on value judgments, while judgments about ordinal preferences could be based on peoples' actual choices. Of, course, this argument excludes even weaker forms comparability (Sen 1982, pp. 264–281). In democratic theory, this position is adopted by Tännsjö (1992, pp. 31–2) and by Riker and Ordeshook (1973, p. 112). It also seems to be Arrow's own position. For this reason, he has been labelled as a "positivist" by some authors (Harsanyi 1979, p. 302). Lehtinen (2007) remarks that, in voting contexts, when there are more than two alternatives ordinal preferences are no more observable than cardinal preferences. If choices have strategic aspects, even ordinal preferences cannot be inferred directly from voters' observable choices. A strictly verificationist criterion of meaning may rule out even judgments about ordinal preferences as "meaningless". And, as a general philosophical programme, verificationism seems to be out of business in any case.

Second, some others see interpersonal comparisons as *ethically irrelevant* even if they were available. According to Schwartz,

> There are worthier and more likely purposes served by instituting collective-choice processes than satisfying participants' preferences to the greatest possible degree: such purposes are to distribute power widely, minimizing the abuse of power, to broaden the pool of ideas by which choices are informed, to enhance people's sense of participation in institutions, and to institutionalize orderly shifting of power. To favor people with intense preferences is to favor people who are bigoted, greedy, meddlesome, etc. (Schwartz 1986, pp. 30–1; cf. also Rawls 1971, pp. 30–1, 361; Saward 1998, p. 78).

The validity of the utilitarian principle—"satisfying preferences to the greatest possible degree"—is disputable. But, as we saw, the intensity comparisons are not only in the interest of the maximizing utilitarian. For example, most principles of fairness presuppose interpersonally applicable measures of satisfaction which go beyond ordinal comparisons.

Third, some theorists think that interpersonal intensity comparisons are useless in democratic theory, as—although they may be conceptually meaningful—there is no effective and ethically acceptable way to make the comparisons needed in collective decisions. If the first and the second criticisms could be ignored, a utilitarian theory of social good or welfare would, in principle, make sense (for a defence of an essentially Benthamite system, see Ng 1979; for a sophisticated Millian alternative, see Riley 1988). But the problem of creating an *institutional* system to collect the necessary information would remain. What is needed is an institutional method of making the required intensity comparisons *en masse*—it would not be helpful if such comparisons could be made, say, in laboratory conditions or by using "extended sympathy" in personal interaction (MacKay 1980, pp. 73–6). Moreover, even if there were an effective way of making interpersonal intensity comparisons, any such method would necessarily be undemocratic. The most plausible conception of democracy contains at least the following normative components: the voters' voting power is (roughly) equal; their choices determine (directly or indirectly) the outcomes; and the choices are free, not coerced or manipulated. We may have different ways of arguing that a million spent on the health care of poor children is, in terms of justice or human welfare or happiness, better used than a million spent on tax cuts for wealthy people. Public organizations, such as welfare agencies, do make such comparisons, and in making them, they may use scientific information as well as everyday knowledge, empathy and imagination. But the information they use is not inferred from valuations consciously given by citizens, nor are they aggregated by using a method that would ensure procedural equality between the respondents.

Roughly, many normative theorists of democracy see the intensity problem as irrelevant for the second reason, and many empirically oriented political scientists see it as relevant but irresolvable for the third reason, while many theorists of public choice and of social choice see the problem both as relevant and solvable. The obvious response to the third critique would be to construct a democratic method which could make systematic intensity comparisons possible. The rest of this chapter is mainly about the most popular proposal.

3 The Borda Rule and Intense Minorities

When a choice is made between two alternatives, majority rule satisfies Arrow's independence condition. Moreover, as Kenneth May has shown in his classical article, majority rule is the only rule which satisfies the further conditions of decisiveness (Arrow's "universal domain"), anonymity (which implies Arrow's "non-dictatorship"), and strong responsiveness (which implies the Pareto condition). However, this positive result cannot be extended to cases with more than two alternatives. If there are more than two alternatives and none of them is the most-preferred alternative for more than a half of the voters, there are several options. We may either drop the "more than half" requirement and be satisfied with mere plurality, or drop "the most-preferred" requirement and try to reduce the choice to a series of pair-wise majority comparisons. The latter is the basis of the well-known Condorcet criterion. For many theorists, the Condorcet criterion is the most plausible extension of the majority principle in voting contexts, or even the only criterion compatible with democracy. Iain McLean makes the argument explicit:

> What is so special about a Condorcet winner? Let us go two steps backwards. What is democracy? Majority rule. Majority rule is necessary, though doubtless not sufficient, to any definition of democracy. What is majority rule? The rule that the vote of each voter counts for one and only one; and that the option which wins a majority is chosen and acted on. Indeed, the second requirement is little more than a special case of the first. For if an option which is not a majority winner is chosen, then the votes of those who supported it turn out to have counted for more than the votes of those who would have supported the majority winner. And that is exactly what happens when a Condorcet winner exists but is not chosen (McLean 1991, p. 177).

Although Condorcet-effective rules do not satisfy Arrow's independence condition, they satisfy it more often than other weakly neutral and anonymous rules, for they are *bound* to violate it only in the cyclical cases. This follows from their basic logic: they reduce complex choices to a series of pair-wise majority choices. Indeed, Michael Dummett (1984)—who does not himself unqualifiedly support the Condorcet criterion—thinks that anyone who sincerely adheres to the absolute-majority principle in dichotomous choice-situations must also adhere to Condorcet's principle when there are more alternatives than two. What really matters for a majoritarian is the number of people satisfied with the result, not the relative *degrees* of satisfaction.

According to Dummett, however, the number of satisfied voters cannot be relevant as such. Ultimately, even the majority principle derives its normative force from "total satisfactions". As he says.

> The question turns on whether it be thought more important to please as many people as possible or to please everyone collectively as much as possible. The latter is surely more reasonable. The rule to do as the majority wishes does not appear to have any better justification as a rough-and-ready test for what will secure the maximum total satisfaction: to accord it greater importance is to fall victim to the mystique of the majority (Dummett 1984, p. 142).

In this interpretation, all voting rules are imperfect measures of the maximum total satisfaction. However, majority rule is not a particularly good measure of total satisfaction unless we have reasons to believe that the intensities are equal (cf. Riley 1990). To make the matter more clear, let us consider the following case:

Example 1

51	49	voters
a	*b*	
b	*c*	
d	*d*	
c	*a*	

In the example *a* is the majority winner, and therefore a Condorcet winner too. One might, however, argue that there would be a good case for selecting *b* instead of *a*. Although a slight majority favours *a*, for a large minority *a* is the worst alternative, while *b* does not offend anyone. It is possible that, by selecting *b* instead of *a*, we may increase the "total satisfaction". Various point-counting rules, of which the Borda count is the best known, would select *b*. If the voters are allowed to give three points for their favourite, two points for their second choice, etc., *b* would receive a total of 249 points against *a*'s 153 points. In the example, *b* is the Borda winner. The Borda count seems to be the most promising way to institutionalize intensity comparisons in voting contexts. It is the rule which has enjoyed continuous support of the specialists since Nicolaus Cusanus[3], and one which has also applied in practice. Plurality, Condorcet, and Borda are commonly conceived as being the three main competing criteria for democratic decision-rules (see, for example, Budge 2000). It may be argued that *all* the other electoral principles are either imperfect substitutes of, or compromises between, these three principles.

Example 1 also shows how intensity considerations may be justified in terms of fairness (rather than in terms of "total satisfaction"). Suppose that we want to avoid "majority tyranny" (or, less dramatically, the problem of "permanent minorities"; cf. Jones 1988; Karvonen 2004) by giving the minorities some real power over the outcomes. If any minority smaller than a half of voters had the power to determine some outcomes, the system would be indecisive, for obviously there could be more

[3] The Borda rule was already described by the great philosopher and theologician Nicolaus Cusanus in his *Concordantia catholica* (1433/1995) where it was recommended as the best method to choose the Emperor. According to Antony Black (1994, p. 39), the Council of Basle (1431–49), in which Cusanus was a member, actually used a Borda-like rule. In Belgium, a Borda-like preferential rule was used in clerical elections from the sixteenth to the eighteenth century, although the weights were not same as in Borda: one first-preference vote was worth of two second-preference votes or three third-preference votes (Moulin 1958, p. 547). The interpretation of the electoral results caused some disagreements. Pope Gregory XV (pope 1621–1623) decided that when "the number of votes" and "the number of voters" pointed in different directions, the latter was decisive (ibid., p. 517). This is the first recorded conflict between these two criteria of preference-aggregation.

than one minority making the claim at the same time. If only some nameable minorities had the power, the resulting quasi-corporativist rule would violate anonymity. Finally, a general minority-veto would favour conservative minorities. In contrast, an intensity measuring rule like the Borda count would give more power to the minorities without violating the requirement of voter equality. For example, with four alternatives, the Borda count guarantees that a majority cannot dictate the outcome in all possible choice-situations, unless it is larger than three-fourths (Nurmi 2007, pp. 116–7). A comparison with approval voting—which is sometimes considered as a "utilitarian" rule (see Hillinger 2005)—is illustrative. Approval voting allows that a narrow majority can guarantee the selection of its favoured outcome under sincere or coordinated strategic voting. (Baharad and Nitzan 2005). Consider Example 1 again. If the 51 voters strategically approve only the alternative *a*, it is selected in spite of the strong and intense opposition. This problem can be mitigated by requiring that the voters should vote (at least) for two alternatives. But this solution would make the rule less sensitive to intensity considerations. Even voters who sincerely reject all but one alternative would be forced to give an equally weighty vote for some of the rejected alternatives.

Because the Borda count possesses the relatively rare *strong* responsiveness property, it guarantees that *all* changes in voters' preference orderings are reflected by the final choice. For this reason, the results of the Borda count actually agree with the Condorcet-criterion more often than the results produced by other positional or semi-positional rules in general use. Thus, the notion of a Borda winner may look like an attractive alternative to the Condorcet criterion. It partly agrees with our majoritarian intuitions while leaving some room for other considerations.

However, these results do not show that the Borda rule actually provides a practicable way to measure *intensity differences*. In his book *Voting Procedures* (1984), Michael Dummett recognizes that many arguments for and against various voting rules are based on suppositions about the typical preference structures. He criticizes the plurality criterion because it looks only at the first preferences. As he remarks, one ground upon which it can be defended is the supposition

> that the gap in any voter's preference scale between any outcome other than his first choice and the next outcome on his scale is not merely small, but infinitesimal, in comparison with the gap between his first choice and his second (p. 132).

This supposition concerns intensity differences, and although it may hold in some cases, it is just one possibility among many. To see Dummett's point, consider a case in which the plurality rule is used to produce a full ranking rather than just choosing the best alternative:

Example 2

99	1	voters
a	*c*	
b	*b*	
c	*a*	

According to the plurality criterion, *c* is the second-best alternative, for *c*, unlike *b*, appears as the first in the preferences of at least one voter. Nevertheless, *all* voters except one rank it lower than *b*; the plurality ranking looks acceptable only if the voters put no weight on their lower preferences. "Certain gaps", says Dummett, "between consecutive outcomes on an individual voter's preference scale may be small, others large; but there can be no general rule for determining which". This is plausible; there seems to be no universal reason why voters themselves would put all the weight on their first preferences. In some cases the distance between the best and the second best may be negligible. Dummett's general conclusion, however, is less plausible: "the only general rule we can reasonably adopt is that all the gaps are not merely comparable, but equal" (p. 133). This sounds like an application of the Principle of Insufficient Reason. Dummett's argument seems to be this: if we do not know what the actual differences are, we have to treat them as equal. But the principle itself is a problematic one. Consider the following possibility: The 51 voters in Example 1 above are actually almost indifferent as between alternatives *b, d* and *c*, but they all agree that these alternatives are much worse than *a.* To make the case more dramatic, let us suppose that the consequences of all the other alternatives than *a* would be perceived as catastrophic by the 51 voters. The 49 voters who favour *b* have no intense preferences over the issue. They could almost as well accept some other result. The measured 'intensities' are, in this case, products of the instrument of measurement; the plurality rule would measure them more accurately. The problem is that *all* such general suppositions, including Dummett's equal distance supposition, are necessarily *ad hoc*. According to one early proposal, voters might give one vote for their favorite and a half vote for the second-best (Dabagh 1934). As a sort of compromise between the plurality and the Borda rule, Dummett (1997, 167–73) recommends a modified Borda rule which awards six points to a party standing highest in a voters' ranking, two points to the second highest preference, and one to the third. However, there are infinitely many ways to assign the weights. Without a general argument, the problem of social preferences has not been solved but only thrust back onto the choice of weights (Feldman 1980, p. 194) Indeed, Sugden (1981, p. 143) admits that his intensity-based argument does not pick Borda as the *uniquely* best "neo-utilitarian" rule.

4 Saari's Argument and the Interaction Problem

In spite of the problem presented above, many defenders of Borda, including Dummett (1984), Saari (1995) and Sugden (1981, p. 144) see it essentially as an imperfect but practicable intensity-measuring device. Saari has, however, provided an extremely interesting argument which is, as such, independent of the intensity considerations. Here, I try to present a short sketch of the basic argument. Consider, first, the following situation

Example 3

5	3	voters
a	*b*	
b	*a*	
c	*c*	

Here, *a* is both the Borda and the Condorcet winner. Now, let us add nine new voters whose preferences exhibit the familiar Condorcet paradox:

Example 4

5	3	3	3	3	voters
a	*b*	*b*	*a*	*c*	
b	*a*	*a*	*c*	*b*	
c	*c*	*c*	*b*	*a*	

According to Saari, these nine additional voters are tied; hence their votes should not be able to change the initial outcome. (Analogously, if we add three voters who prefer *a* to *b* and three with the opposite preference, this group of six is tied, and should not change the outcome!) However, in Example 4, *b* becomes the Condorcet winner. Alternative *b* beats alternative *a* 9–8 and alternative *c* 11–6. In contrast *a* remains as the Borda winner in both examples, even after the invasion of the nine new voters. Their votes—three first places, three second places and three-third places for each alternative—cancel out each other. According to Saari, *this phenomenon accounts the whole Arrowian indeterminacy problem.*

Saari formulates two symmetry requirements:

The *Neutral Reversal Requirement*: When two rankings reverse one another, say $a > b > c$ and $c > b > a$, they are tied and do not change the outcome.

The *Neutral Condorcet Requirement*: When *n* rankings over *n* alternatives form a complete cycle, say $a > b > c$, $b > c > a$ and $c > a > b$, they are tied and do not change the outcome.

Majority rule respects the Neutral Reversal Requirement but not the Neutral Condorcet Requirement. In contrast, all positional rules (including the plurality and the Borda rules) respect the Neutral Condorcet Requirement, but *only* the Borda rule also respects the Neutral Reversal Requirement. Thus, the Borda rule is the best voting rule. According to Saari, this conclusion can be challenged only by showing that the Neutral Condorcet Requirement is not relevant, in other words, that a symmetric cycle between alternatives should *not* be treated as a tie.

The real defect of the Condorcet criterion is that pair-wise comparisons mandated by the independence condition do, according to Saari, disregard some important information about the preferences of the voters. Consider a voting cycle: *a* defeats *b, b* defeats *c* and *c* defeats *a* in a series of majority contests. This may result from an underlying Condorcetian cycle of majority preferences. But it might also result from intransitive *individual* preferences: some voters have simply voted in an irrational way. We cannot tell the source of intransitivity by looking at the pair-wise voting results. Saari's point is not that such a situation is likely to occur, or that a voting rule should be able to deal with it; the point is that a good rule should be able to distinguish between the two sources of intransitivity. By excluding all information not related to the ordinal preference rankings, Arrow's independence condition also excludes essential information about the nature of these rankings. As Saari puts it, "losing the intensity information corresponds to dropping the critical assumption that voters have transitive preferences". While the Borda rule does not satisfy Arrow's independence condition, it is the only rule that satisfies the *binary intensity independence* condition which requires that the relative ranking of each pair of alternatives be determined by voters' relative rankings of that pair, and that the intensity of this ranking is determined by the number of candidates ranked between them (Saari 1995, pp. 201–2).

Saari's writings are not only mathematically innovative but also philosophically sophisticated. He sees the Arrow theorem as one instance of a general problem of information aggregation, and finds analogical problems in sports, statistics, law, engineering, and economics. All his examples illustrate the problems which appear when we try to understand or evaluate a whole by aggregating information achieved from its parts. He warns: "Expect paradoxical phenomena whenever there is a potential discrepancy between the actual unified whole and the various ways to interpret the totality of disconnected parts" (Saari 2001, p. 104). The great merit of Saari's approach is that several apparently unrelated but somehow "paradoxical-looking" phenomena are shown to be instances of a single general problem. It does not follow, however, that there exists a corresponding single *solution*, applicable in all contexts. My thesis is that voting in political contexts has specific properties which are not present in the other cases discussed by Saari.

Consider the following example:

Example 5

3	2	2	voters
a	*b*	*c*	
b	*c*	*a*	
c	*a*	*b*	

In this example the introduction of a Pareto-dominated alternative c^* reverses the ordering of alternatives based on the Borda-criterion. Without it, a gets 8, b gets 7 and c gets 6 points When it is introduced, the Borda scores are: 6 for c^*, 11 for a, 12 for b and 13 for c.

3	2	2	voters
a	*b*	*c*	
b	*c*	*c**	
c	*c**	*a*	
*c**	*a*	*b*	

Other preference counting rules—STV, the Bucklin rule, and the supplementary vote—produce similar if somewhat less dramatic anomalies. (On this anomaly in STV, see Doron 1979). These effects cannot plausibly be interpreted in terms of intensity differences. Suppose, for example, that c^* is in all essential aspects identical with c, but contains some technical defect and is therefore considered worse than c by all the voters. For an informed decision, its presence on the agenda is totally irrelevant, for it does not contain any new aspect not already contained in c.

Example 5 shows that the *agenda-setting process* is crucial for the Borda rule. The problem presented above is, of course, well known by the proponents of the rule. Some of them (for example, Dummett) have argued that agenda manipulation is less likely to cause troubles in real elections, for it may be difficult to produce suitable "dummy" alternatives (like c^* in the example above). Mackie (2003, pp. 153–155) claims that someone who tries to manipulate a voting rule by addition or subtraction of alternatives needs to know voters' exact preference rankings, including their rankings over manipulative alternatives (like c*).

In order to assess these empirical claims, let us consider the almost only example of the use of the Borda rule in politically important decision-making: the choice the candidates for the office of *Beretitenti* or the president in the island-state Kiribati. According to the constitution of Kiribati, the legislature (*Maneaba*) chooses three or four candidates; and one of them is elected by the people to the office. The candidates are selected in *Maneaba* by using a limited version of the Borda count. There can be many candidates, but members of *Maneaba* are allowed to rank four of them. Those four having largest scores are allowed to continue in the final (popular) contest. In 1991, there were eight candidates presented for the *Maneaba*. According to Ben Reilly (2002, pp. 367–9), there was extensive strategic voting in which two of the most popular candidates were played out from the final election. Two of the running candidates were "dummies". Their role was exactly the same as that of the alternative c^* in our example: by voting a "dummy" alternative the voters could avoid giving any lower preference support for the most serious challengers of their favourite candidates. In the only politically relevant real-life case described in the literature, the Borda count worked *exactly* as its critics expected it to work. Reilly quotes another commentator of the Kiribati election "It remains to be seen just how long such a system will be tolerated which has the effect of eliminating popular candidates through backroom political manoeuvring" (p. 368).

This form of manipulation is particularly attractive when the Borda count is used. According to Serais (2008, p. 8), in three-candidate Borda elections the a priori probability of situations which can be manipulated by "cloning" alternatives is always over 40 %, and approaches rapidly to 62 % when the number of voters increases. *Pace* Mackie, the manipulators need not to know the exact preference rankings; it is sufficient for their purposes if they can produce alternatives which are generally perceived as 'clones' of their preferred alternative. The resulting multiplication of the Borda scores guarantees that *some* among the essentially similar alternatives will be selected—unless, of course, the other groups are able to use the same strategy. If, for example, the Borda rule were used for allocating seats between parties in an assembly, a party might increase its share of seats by splitting itself up to two essentially similar but nominally different parties. The point is nicely illustrated in Sverker Härd's study on seat allocation rules in the *Riksdag* of Sweden (Härd 1999, 2000). Using opinion measurements, Härd simulated the distribution of seats in the Swedish Parliament under different voting rules. One of the rules tested by Härd was a version of Borda. In this application, a party's proportion of the seats in the *Riksdag* was the same as its proportion of the total amounts of the Borda points. The result was a massive shift of power from the Social Democrats to the small non-Socialist party groups. The obvious reason for this shift—not discussed by Härd—is that in Sweden the non-Socialist party groups are numerous, while in the Left the only alternatives are the Social Democratic party and the small Leftist (ex-Communist) party. The number of ideologically close parties multiplied their compound Borda scores. If the Borda rule were actually used in the Swedish elections, the Left could regain its power simply by creating more, nominally independent groups. A general result proved by van der Hout et al. (2006, pp. 465–7) shows how problems of this type can be avoided only by using first preference information as the sole basis for seat allocation.

There is a further problem. The Borda rule is likely to produce larger set of candidates than, say, the plurality rule. Intuitively, the reason is that candidates who do not have much first-preference support still have some hopes to get elected. Any rule that takes some of the lower preferences into account tends to have this effect, even without any conscious attempts to manipulate the agenda. Ordinary voters are not necessarily able to produce strict and complete preference orderings when the number of alternatives becomes large (say, over five). It is reasonable to expect that voters are generally able to submit transitive preference orderings, as Saari says. It is, however, less obvious that the rankings submitted by them would always satisfy the strictness or completeness requirements. If voters are nevertheless required to submit strict and complete rankings (as in the Australian alternative-vote elections) an elections result may actually be determined by voters who—when unable to rank all the candidates—fill their ballot papers in a random way. Therefore, a reasonable voting system should either to limit the number of candidates, or to allow incomplete ballots.

However, while modified versions of the Borda count can handle incomplete rankings, there are inevitable costs (Nurmi 2007). First, such modifications are vulnerable to strategic truncation of preferences. In many voting situations, it is rational not to submit one's complete preference ordering (for such truncation strategies, see Lagerspetz 2004). Second, all attempts to modify the Borda rule are likely to undo some of the most attractive properties of the rule. Most notably, the modified versions may elect a candidate who is considered as the worst by a majority of voters. Given the effect exemplified in Example 5, these results are to be expected: if the removal of a candidate from the contest—c^* in the example—may change the outcome, his removal from sufficiently many ballot-papers may have a similar effect. If these costs are unacceptable, the remaining solution is to limit the number of alternatives beforehand. While this may reasonable in some contexts—for example, in multi-alternative referendums—in general elections it is clearly incompatible with the principle of democratic freedom.

Thus, there is an important difference between voting and the other aggregation contexts analysed by Saari. Only in the context of voting, the choice of the method of aggregation may change the *input* of aggregation. This reflects a general problem shared by many attempts to "apply" the results of social sciences. In engineering, statistics etc., the reality itself does not react to the choice of method of acquiring information about it[4]. Hence, the manipulative aspects of the Borda rule may well be irrelevant in such contexts (on engineering contexts, see Scott and Zivikovic 2003). In contrast, voters' strategies, the composition of agendas, the supply of candidates etc. may vary with the chosen voting rule. This adds to voting situations an additional element of arbitrariness not present in Saari's other examples. The question is not just which method would reflect the objects (voters' opinions) in the most accurate way, but rather, which would be the best method *given the unavoidable interaction between the aggregation process and the objects of aggregation*.

Saari's argument for the Borda rule, brilliant as it is, should be balanced against the defects of the Borda rule discussed above. A Condorcet-effective rule is sensitive to the addition of new (tied) voter groups, but, as we saw, the Borda rule is sensitive to the addition of new (Pareto-dominated) alternatives. If the Condorcet criterion loses some information about the transitivity of the rankings, the Borda rule lets in some questionable information. The normative interpretation of the Borda rule is, even for Saari, that it is able to take preference intensities into account. But if the number of candidates between a and b in someone's expressed preference orderings may reflect other factors than preference intensities, it is difficult to argue that this information should have an effect on the final choice. While Arrow's independence condition is too strong (it may leave out some relevant information), Saari's alternative condition is too weak (it allows that irrelevant information may determine the outcome). Personally, I am unable to decide which form of arbitrariness disturbs me more.

[4] In an early paper (Lagerspetz 1988) I argued that the main methodological difference between the natural and the social sciences is related to this possibility.

5 For and by the People

In many works informed by the theory of social choice, the underlying supposition is that the main purpose of voting rules is to aggregate *information*. A voting rule is, indeed, a means of aggregation. The fundamental issue is how the results of aggregation are to be interpreted. We may distinguish two different ways to interpret voting, and, correspondingly, two partly different perspectives from which a voting rule may be evaluated. According to one view, the task of the voting rule is the provide information about some independently existing properties of the world, basically of voters' preferences. Thus, voting is a kind of *measurement*, and the aggregation problems appearing in political contexts are largely analogous to those appearing in statistics, multi-criteria decision-making etc. A voting rule should be as reliable and exact instrument of measurement as possible. For example, Claude Hillinger (2004) compares voting to measurements in sociology, psychology, market research etc., and remarks that in these contexts cardinal scales are always used. "It is only in voting and particularly in political voting, that the scales are restricted. For this there is no apparent reason, nor, as far as I know, has any argument in defence of this practice been advanced". Thus Hillinger (2004, 2005), like Ian Budge (1996, pp. 164–5), argues for cardinal scoring rules. Budge defends his proposal with the same analogy: "similar procedures are used in psychological tests and opinion polls with results which are widely accepted" (p. 165). He comments on the possibility of strategic behaviour: "Voters in the mass are also likely to assign scores that reflect their true feelings, unless urged to engage strategic misrepresentation by political parties. *But these can, if necessary be legally forbidden to do so*" (*idem*, emphasis EL). The last sentence reveals one difficulty in the measurement interpretation of voting. Is it compatible with democratic freedom that people—with or without party affiliations—are not allowed to give voting recommendations to their fellow citizens?

The problem of strategic behaviour reveals an interesting difference between voting and measurement. As Sager (2002, p. 185) remarks, strategic behaviour *may* be a problem even in social measurement if the subjects expect that the results are utilized in decision-making. Consequently, questionnaires are often designed in a way that makes it hard for informants to see how their answers can influence future policy decisions. In voting contexts, the democratic ideal requires that the connection between the answers given and the future policy decisions *is* as clear as possible. Indeed, various institutions (for example, proportional representation, coalition governments, bicameralism, representative institutions in general) are often criticized for the *lack* of a visible connection between votes and future policy decisions.

A further argument against the measurement interpretation is that it does not provide any justification for democratic equality. Suppose that, in order to save election costs, we select 1/10 of adult population as the *demos.* Only those belonging to this selected group are entitled to participate in referenda or in general elections. If we use the modern techniques of random sampling in choosing the *demos*, the distribution of opinions and interests in the demos will mirror the general population very accurately. Consider normal opinion measurements. By using small random samples (*much* smaller than 1/10 of the electorate), the pollsters are able to predict the choices of the total population with a great degree of precision. With an enormous sample of 10 % of the total population, the deviation would be negligible. The randomly composed *demos* would elect the same candidates and vote for the same parties in equal proportions as the entire population. If the main purpose of voting were to provide information, recording *everybody's* preferences seems to be just a waste of time and money[5].

There is, however, another possible interpretation of voting. It should not be seen mainly as a means to get information. It is primarily an *exercise of power*. To take the obvious case, when voting in a parliament, the MP's are not providing information *about* their opinions. They are making binding decisions *based on* those opinions. Elections can be interpreted in the same way. It is, of course, plausible to say that an elections result usually provides information, mostly about the relative popularity of parties and candidates but also about other issues (for example, the turnout rates may measure political alienation). The *main purpose* of elections, however, is not to provide information but to choose the most popular candidates. A good voting rule should produce outcomes which are recognized as legitimate. In order to produce legitimate results the rule must be compatible with the background values; in democracies these values include equality, liberty, and effective voter influence. Because voting is also an exercise of power, voters are—and should be—moved by motives which are not operative when the same people are filling in questionnaires or answering questions in an opinion poll. As Saward (1998, p. 35) says, an opinion poll can gather expressions of preference, but they are not preferences which reflect the fact that people are aware that their expressions will decide anything[6]. Because of the power aspect, elections are taken seriously; and this unavoidably provides incentives *both* for rational deliberation *and* for strategic behaviour. This does not, however, mean that there are no normative problems related to strategic behaviour in democratic elections. The social choice results tell us that strategic manipulation is possible in all democratic systems. A realistic aim is to minimize the role of certain forms of strategic behaviour.

One possible counterargument[7] to my analysis is this. In mass elections the probability that an individual voter would be decisive is extremely small. If power is measured in terms of decisiveness, the power exercised by an individual voter is almost zero. This creates a collective action problem. A candidate may win only if

[5] Fishkin's "deliberative poll" (Fiskin 1991) operates with a randomly selected *demos*. However, although he recommends its use as an aid in democratic decision-making, he does not propose that it should replace general elections.

[6] Fishkin (1991, p. 83) quotes a study on opinion measurements: "Most respondents feel obliged to have an opinion, in effect, to help the interviewer out. (…) In effect, opinions are invented on the spot".

[7] This counterargument was made by Manfred J. Holler.

a sufficient number of citizens' vote for him or her; and, more generally, democracy can work and produce legitimate decisions if sufficiently many citizens are willing to participate. But nevertheless, a single citizen has no convincing instrumental reason to cast an informed vote, for his or her personal contribution to the outcome is likely to be negligible in any case. Perhaps voting should be interpreted as a purely expressive act like cheering in a soccer match (Brennan and Lomasky 1993).

This argument certainly points out a real problem (first discussed by G. W. F. Hegel in his famous article on the Estates of Württemberg) for the view that voting acts could be interpreted as purposive exercises of power. However, it does not work as an argument for the measurement view of voting. If voting is an expression of feelings, voting results do not measure voters' preferences over outcomes in a reliable way. The expressive interpretation implies that voters are actually choosing between alternative *voting acts* ("How do I feel if I cast my vote in this way ?") rather than between competing candidates or policies.

There is not enough space for a convincing answer, but some observations can be made. A purely expressive model cannot explain the fact that when there are competing acceptable candidates or parties, people are more willing to give their support to those that have realistic chances to succeed, given the expected choices of the others. In practice, people tend to vote for an acceptable candidate who has realistic prospects to be elected, rather than for the candidate who might be their absolute favourite. *All* electoral systems tend to constrain political competition as a contest between a limited number of realistic candidates or parties. The most plausible explanation of this (Cox 1997) appeals to voters' instrumental rationality. But we have already admitted that instrumental rationality cannot explain why people vote at all! A solution of this dilemma is, I think, that a voting act is (at least sometimes) seen as a contribution to a collective action. In mass elections, voters are (at least sometimes) motivated by a "consequentialist generalization". In other words, they ask themselves: "What would happen if *all* (or most, or very many) *people like me* would choose in this way ?" Voters tend to portrait themselves as participants in collective actions. They try to evaluate the consequences of those actions rather than the consequences of their individual voting acts.

There is a more general philosophical lesson in the distinction between measurement and voting. Real-life rules of social choice do not connect voter's preferences directly to outcomes. Instead, they connect *expressions* of preferences—votes to outcomes. Suppose that we had a measurement device that would connect (ordinal or cardinal) preferences directly to outcomes, say, by measuring peoples' neural states. Suppose, moreover, that the officials—a benevolent autocrat or central planning agency—would then implement the outcomes that were picked by the aggregated measurement results. Would that constitute a democratic arrangement? The answer is, I think, no. Why? In the thought example, there would be no element of popular *choice* or *authorization* by the citizens. The citizens' role would be a purely passive one. The system would constitute a government *for* the people, not *by* the people. It would give people what they desired, not what they would have desired *when knowing that a public expression* *of their desires causally contributes to, and therefore makes them responsible for, the resulting outcomes*. These are likely to be different things: the authoritative nature of the voting process forces voters to consider their preferences and the way their votes are connected to the outcomes.

6 Conclusion

Voting rules are used for different purposes. Votes are taken in representative bodies, general elections and referendums, as well as in multi-member courts, panels of experts, collegial bodies, and public contests. The rules cannot be evaluated without taking wider institutional and social contexts into account. More specifically, it is not possible to find “the best” rule simply by comparing the performance of various voting rules in respect with the pre-given criteria of social choice (Lagerspetz 2004, pp. 218–20). When, for example, we have the luxury of choosing between the Borda rule and some Condorcet–effective procedure, we should appeal to pragmatic and context-dependent considerations for and against both alternatives. In many contexts, the Borda rule may be preferable. If the set of alternatives is fixed, the effects discussed above cannot occur. For example, when we are pooling experts’ judgments or the popular judgments on the performance of competing contestants (for example, in the Eurovision song contests), the “agendas” are exogenously given. The Borda rule may well be the most plausible method to aggregate information in such contexts. But in such contexts, “intensities”, conceived in the utilitarian way, are not relevant. If we decide to use the Borda rule, our reasons should *not* be related to intensity considerations.

To quote Sartori (1987, p. 225) “the intensity criterion cannot establish a workable rule”. However, democratic processes are not insensitive to varying intensities of preference. Pressure group activities (Dahl 1956), vote trading (Buchanan and Tullock 1962), decentralized decision-making (Karvonen 2004) and public argumentation can all be seen as informal ways to cope with the intensity problem. They are necessarily unsystematic, imperfect and partial solutions. But, given the nature of the problem, this necessity may actually be a virtue.

Literature

Baharad, E., & Nitzan, S. (2005). Approval voting reconsidered. *Economic Theory*, *26*, 619–628.

Black, A. (1994). Diplomacy, doctrine and the disintegration. In J. Helmrath & H. Müller (Eds.), *Studien zum 15. Jahrhundert: Festschrift für Erich Meuthen*. München: R. Oldenbourg.

Brennan, G., & Lomasky, L. (1993). *Democracy and decision. The pure theory of electoral preference*. Cambridge: Cambridge University Press.

Buchanan, J. M., & Tullock, G. (1962). *The Calculus of Consent*. Ann Arbor: University of Michigan Press.

Budge, I. (1996). *The new challenge of direct democracy*. Cambridge: Polity Press.

Budge, I. (2000). Deliberative democracy versus direct democracy—plus political parties! In M. Saward (Ed.), *Democratic innovation deliberation, representation and association*. London: Routledge.

Christiano, T. (1996). *The rule by the many*. Boulder: Westwiew Press.

Cox, G. (1997). *Making votes count: Strategic coordination in the world's electoral systems*. Cambridge: Cambridge University Press.

Dabagh, T. S. (1934). The vote-and-a-half system of majority choice by plurality rule. *The American Political Science Review, 28*, 93–101.

Dahl, R. A. (1956). *A preface to democratic theory*. Chicago: University of Chicago Press.

Doron, G. (1979). Is the hare voting scheme representative ? *The Journal of Politics, 41*, 918–922.

Dummett, M. (1984). *Voting procedures*. Oxford: Oxford University Press.

Dummett, M. (1997). *The principles of electoral reform*. Oxford: Oxford University Press.

Feldman, A. M. (1980). *Welfare economics and social choice*. Boston: Martinus Nijhoff.

Fishkin, J. S. (1991). *Democracy and deliberation. New directions for democratic reform*. New Haven: Yale University Press.

Härd, S. (1999). *Den godtyckliga demokratin: En studie av olika metoder att tilgodse kravet om proportionalitet*. Uppsala: Uppsala University Library.

Härd, S. (2000). Arbitrary democracy. In N. Berggren, N. Karlson, & J. Nergelius (Eds.), *Why constitutions matter*. Stockholm: City University Press.

Harsanyi, J. (1979). Bayesian decision theory, rule utilitarianism and Arrow's impossibility theorem. *Theory and Decision, 11*, 280–317.

Hillinger, C. (2004). On the possibility of democracy and rational collective choice, http://epub.ub.uni-muenchen.de/429/1/DEMOCRACY.pdf.

Hillinger, C. (2005). The case for utilitarian voting. *Homo Oeconomicus, 22*, 295–321.

Jones, P. (1988). Intense preferences, strong beliefs and democratic decision-making. *Political Studies, 36*, 7–29.

Karvonen, L. (2004). The Achilles' heel of democracy? on majority and intensity. In M. Wiberg (Ed.), *Reasoned choices* (pp. 242–263). Turku: The Finnish Political Science Association.

Kendall, W., & Carey, G. W. (1968). The intensity problem and democratic theory. *The American Political Science Review, 62*, 5–24.

Lagerspetz, E. (1988). Reflexive predictions and strategic actions. *Social Science Information, 27*, 307–320.

Lagerspetz, E. (2004). Evaluating voting systems. In M. Wiberg (Ed.), *Reasoned choices* (pp. 201–222). Turku: The Finnish Political Science Association.

Lehtinen, A. (2007). A farewell to Arrow's theorem. http://philsci-archive.pitt.edu/archive/00003429/01/fiia.pdf.

Mackie, G. (2003). *Democracy defended*. Cambridge: Cambridge University Press.

MacKay, A. (1980). *Arrow's theorem. The paradox of social choice, a case study in the philosophy of economics,* London: Yale University Press.

McLean, I. (1991). Forms of representation and systems of voting. In D. Held (Ed.), *Political theory today* (pp. 172–196). Cambridge: Polity Press.

Moulin, L. (1958), "*Sanior et maior pars*. Note sur l'evolution des techniques électorales dans les ordres religieux du VIe au XIIIe siècle," *Revue historique de droit Français et étranger* 35:368–397, 491–529.

Nicholaus of Cusa [Nicolaus Cusanus] (1434/1995), *The Catholic concordance*, ed. P. E. Sigmund, Cambridge: Cambridge University Press.

Ng, Y.-K. (1979). *Welfare economics. Introduction and development of basic concepts*. London: Macmillan.

Nurmi, H. (2007). Assessing Borda's rule and its modifications. In P. Emerson (Ed.), *Designing an all-inclusive democracy. Consensual voting procedures for use in parliaments, councils and committees* (pp. 109–119). Berlin: Springer.

Rawls, J. (1971). *A theory of justice*. Oxford: Oxford University Press.

Reilly, B. (2002). Social choice in the south seas: Electoral innovation and the Borda count in the Pacific Island Countries. *International Political Science Review, 23*, 355–372.

Riley, J. (1988). *Liberal Utilitarianism. Social choice and J.S. Mill's philosophy*, Cambridge: Cambridge University Press.

Riley, J. (1990). Utilitarian ethics and democratic government. *Ethics, 100*, 335–348.

Riker, W. H., & Ordeshook, P. C. (1973). *An introduction to positive political theory*. Englewood Cliffs: Prentice Hall.

Saari, D. G. (1995). Inner consistency or not inner consistency: A reformulation is the answer. In W. A. Barnett, H. Moulin, M. Salles, & N. Schofield (Eds.), *Social choice, welfare and ethics* (pp. 187–212). Cambridge: Cambridge University Press.

Saari, D. G. (2001). *Decisions and elections. Explaining the unexpected*. Cambridge: Cambridge University Press.

Sager, T. (2002). *Democratic planning and social choice dilemmas. Prelude to institutional planning theory*. Aldershot: Ashgate.

Sartori, G. (1987). *The theory of democracy revisited. Part One: The contemporary debate*. New Jersey: Chatham House Publishers.

Saward, M. (1998). *The terms of democracy*. Cambridge: Polity Press.

Schwartz, T. (1986). *The logic of collective choice*. New York: Columbia University Press.

Scott, M. J. and Zivikovic, I (2003). "On Rank Reversals in the Borda Count," *Proceedings of DETC 2003*, Chicago: The 2003 ASME Design Engineering Technical Conferences. http://design.me.uic.edu/mjscott/papers/dtm48674.pdf.

Sen, A. K. (1982). Interpersonal Comparisons of Welfare. In A. K. Sen (ed.), *Choice, Welfare and Measurement*, Oxford: Blackwell, 264–281.

Serais, J. (2008). Cloning manipulation of the Borda rule. http://www.hss.caltech.edu/Events/SCW/papers/seraj.pdf.

Sugden, R. (1981). *The political economy of public choice: An introduction to welfare economics*. Oxford: Martin Robertson.

Tännsjö, T. (1992). *Populist democracy: A defence*. London: Routledge.

Van det Hout, E., de Swart, H., & ter Veer, A. (2006). Characteristic properties of list proportional representation systems. *Social Choice and Welfare, 27*, 459–475.

List Apparentements in Local Elections: A Lottery

Friedrich Pukelsheim and Peter Leutgäb

1 List Apparentements

List apparentements form a peculiarity of certain proportional representation systems. In some countries they are employed at the national level, as in Switzerland and Israel. In Germany they are restricted to the local level. Here we elucidate their role in a case study, the 2008 local elections in the German State of Bavaria. Bochsler (2009) presents a more general overview of the subject.

Political parties, or groups of citizens who submit a list of candidates, may register a *list apparentement*[1] with the electoral bureau prior to the election. On Election Day, the conversion of votes into seats then takes place in two stages. Firstly, in the *super-apportionment*, the votes cast for the partners of the apparentement are totaled, and this total enters as a single count into the calculation to apportion all available seats.

This paper has been published in *Essays in Honor of Hannu Nurmi*, Volume I (Homo Oeconomicus 26: 489–500), 2009. We would like to thank M. Holler, Hamburg, and B. Torsney, Glasgow, for valuable comments. The authors' articles quoted in the sequel may be retrieved from the Internet at www.uni-augsburg.de/pukelsheim/publikationen.html. A German version of the present paper has been published in *Stadtforschung und Statistik*, 2/2009.

[1] The French term "apparentement" is also used in English, see Gallagher and Mitchell (2005), p. 631.

F. Pukelsheim (✉) · P. Leutgäb
Institute for Mathematics, University of Augsburg, 86135 Augsburg, Germany
e-mail: Pukelsheim@Math.Uni-Augsburg.De

P. Leutgäb
e-mail: Peter.Leutgaeb@GMX.Net

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*, DOI: 10.1007/978-3-642-35929-3_7, © Springer-Verlag Berlin Heidelberg 2013

Table 1 List apparentements in the local electionsof the State of Bavaria, 2 March 2008

Number of lists	3	4	5	6	7	8	9	10	11	12	13	Sum
One apparentement	72	141	116	64	40	17	5	1				456
Two apparentements		13	45	54	30	27	13	6	2	1		191
Three apparentements				3	4	6	3	4			1	21
Number of communities	72	154	161	121	74	50	21	11	2	1	1	668

Secondly, in the *sub-apportionment*, every list apparentement undergoes a follow-up calculation. Here the seats that the apparentement earned as a whole are apportioned among its partners, proportionally to the vote count for each partner list.

Apparentements do not commit the partner lists during the upcoming legislative period, neither to strive for common goals, nor to enter into a formal coalition. Any party may team up with any other party. There is an affinity of conservative parties to go along with other conservative parties, of course, and liberal groups with other liberal groups. Yet, in our Bavarian case study, we could not identify a definite pattern of who joins which apparentement. Everything is possible, and almost everything is realized.

In the 2008 local elections in Bavaria, just one list apparentement was registered in 456 communities,[2] with the number of campaigning lists running from 3 through 10. Two apparentements emerged in 191 communities, three in 21. Altogether the election featured 901 list apparentements,[3] in 668 out of 2,127 communities. See Table 1.

List apparentements must not be taken as an oath of disclosure towards voters, as is apparent on the ballot paper. Partner lists are not marked in a way that every voter instantly recognizes the affiliation of a party to an apparentement. But seek, and ye shall find. On Bavarian ballot sheets it is the small print, down in the bottom line.

The partners of a list apparentement join companionship only for the day of reckoning. As soon as the electoral results are publicized, the composition of the apparentement disappears from the statistical tables, as if documenting them would constitute an embarrassment to those concerned. What, then, makes list apparentements attractive?

List apparentements are beset with the mystic aura that they even out detrimental disparities of the electoral system. We shall show that such speculations are sometimes right, and sometimes wrong. Moreover, the 2008 Bavarian elections featured thirty-six instances where list apparentements grotesquely reversed the popular vote, in that of two lists the weaker list won more seats.

[2] We use the term *community* as a generic synonym for political entities where voters elect a local council, such as cities, counties, townships, villages, and the like, as in Pukelsheim et al. (2009).

[3] We get $456 \times 1 + 191 \times 2 + 21 \times 3 = 901$.

2 Seat Biases

The element of the electoral system that is notorious for its built-in disparity is the formula for the conversion of votes into seats that comes under the names of D′Hondt, Hagenbach-Bischoff, or Jefferson. We prefer to call it the *divisor method with rounding down (D′Hondt)*, in order to indicate how it works. Any vote count is divided by a common divisor, the *electoral key*, and the resulting quotient is rounded down to its integer part to obtain the seat allocation. The value of the electoral key ascertains that all available seats are handed out (Pukelsheim 2002).

The divisor method with rounding down (D′Hondt) is notorious for its seat biases in favor of larger parties and at the expense of smaller parties. On average, larger parties are allocated more seats than strict proportionality would grant them, and these seats are taken away from smaller parties. There are unbiased alternatives which are increasingly taking over, especially in Germany (Pukelsheim 2003). Among them are the quota method with residual fit by largest remainders (Hamilton/Hare) and, as of recently, the divisor method with standard rounding (Webster/Sainte-Laguë).

Historically, the coupling of D′Hondt with list apparentements is the rule and, in German States, prevails in Bavaria and the Saarland. List apparentements are removed from the law as soon as an unbiased electoral formula is implemented provided the law-makers understand their electoral system, as in the Swiss Cantons of Zürich, Schaffhausen, and Aargau (Pukelsheim and Schuhmacher 2004). Otherwise they remain in the law as a relict of times passed (Rhineland-Palatia). And occasionally an electoral law with old ballast is recycled to give democracy a new start (Thuringia).

The notion of *seat bias* designates the mean deviation of the seats practically apportioned, from the ideal share of seats granted by theoretical proportionality. The mean is evaluated uniformly across all conceivable vote outcomes. Surplus and deficit materialize per each election, and stay practically constant over all council sizes. Seat bias formulas for the divisor method with rounding down (D′Hondt) are listed in Table 2.[4]

Without list apparentements, seat biases exhibit a clear trend. The decrease from profits to losses follows the final vote count ranking. The upper third of stronger lists (in terms of votes received) is granted a surplus of seats. But one man’s meat is another man’s poison. The lower two thirds of weaker lists have to endure a seat deficit.

With list apparentements the seat biases do remain calculable. However, the clear order from top to bottom is lost, and a bewildering diversity of results comes to light. The bewilderment is caused by the double application of the divisor method with rounding down (D′Hondt), thus reinforcing its built-in seat biases. Whether a party wins or loses seats turns into a lottery.

[4] Without list apparentements the formulas are derived in Schuster et al. (2003). With list apparentements the formulas are new and due to Leutgäb (2008).

Table 2 Formulas for the D′Hondt seat biases

Without any list apparentement, the D′Hondt seat bias of the *j*-strongest (in terms of votes received) list is

$\mathrm{D'H(j)} = \frac{1}{2}(\ell\, s(j) - 1)$, where $s(j) = \frac{1}{\ell}\left(\frac{1}{j} + \cdots + \frac{1}{\ell}\right)$.

Here $s(j)$ is the expected vote share of the j -strongest of ℓ lists.

With ℓ lists partitioned into the apparentements $L_1, \ldots, L_k$, the D′Hondt seat bias of the *j*-strongest list becomes

$\mathrm{D'H}\,(j\,|\,L_1,\ldots,L_k) = \frac{1}{2}\left(k\,s(j) + (p-1)\,\frac{s(j)}{s(V)} - 1\right)$, where $s(V) = \sum_{i\in V} s(i)$.

Here V is the apparentement in which the *j*-strongest list figures as one of p partners, and $s(V)$ is its expected seat share.

Example: City of Friedberg, 2008. Of six lists, the second-, third- and fifth-strongest lists joined in an apparentement (case B).

	A: 1,2,3,4,5,6	B: 2 + 3+5,1,4,6
List 1	0.725	0.317
List 2	0.225	0.507
List 3	−0.025	0.160
List 4	−0.192	−0.295
List 5	−0.317	−0.245
List 6	−0.416	−0.444

Without list apparentements (case A), the strongest List 1 may expect an advantage of about 3 seats in 4 elections (3/4 ≈ 0.725). However, Lists 2, 3 and 5 formed an apparentement while Lists 1, 4 and 6 stood for themselves (case B). In this constellation, the largest bonus (0.507) goes to List 2. The total bias increases from 0.950 (= 0.725 + 0.225) in case A, to 0.984 (= 0.317 + 0.507 + 0.160) in case B.

The City of Friedberg (AGS[5] 09771130) provides an instructive example. Six lists campaigned which we retrospectively number from 1 to 6 according to their popular support. That is, List 1 finished strongest and won a larger popular vote than List 2. List 2 entered into an apparentement with Lists 3 and 5, while the others stood alone. The apparentement ranked top in the super-apportionment, where it won a rank-1-bonus. In the sub-apportionment the bonus was passed on to List 2 which was strongest among the partners of the apparentement. The arrangement thus secured a top rank for List 2 twice, in the super-apportionment and in the sub-apportionment. In the end, the weaker List 2 won more seats than the stronger List 1. The Bavarian electoral law circumnavigates the popular vote, by way of list apparentements.

[5] AGS = Amtlicher Gemeindeschlüssel = official community key. The key defines a standard order for German communities. It may also be used to retrieve some basic statistical information about the community via www.destatis.de/gv/.

Table 3 D′Hondt seat biases for three-party systems, empirical and theoretical values, Bavarian local elections 2008

List partitions	A: 1, 2, 3		B: 1, 2 + 3		C: 1 + 2, 3		D: 1 + 3, 2	
	Empir.	Theor.	Empir.	Theor.	Empir.	Theor.	Empir.	Theor.
Strongest list	0.218	0.416	0.000	0.111	0.333	0.455	0.266	0.534
Median list	−0.019	−0.083	0.137	0.135	0.000	−0.066	−0.133	−0.222
Weakest list	−0.019	−0.333	−0.137	−0.246	−0.333	−0.389	−0.133	−0.312
Total bias	0.218	0.416	0.137	0.246	0.333	0.455	0.266	0.534

3 Three-Party Systems

The analysis remains somewhat more transparent in three-party systems, the simplest constellation where list apparentements come into play. With only a single list the election turns into a simple majority vote. When there are two lists (2008 in Bavaria in about four hundred communities), they are lacking a third against whom it would pay to join into an apparentement.

Although three-party systems represent the simplest case, it is sufficient to indicate potential complications since there exist four ways of partitioning the lists. In case A (1, 2, 3) all lists stand alone. In the cases B, C und D a two-partner list apparentement is formed. In partition B (1, 2 + 3) the two weaker lists join in an apparentement, in C (1 + 2, 3) the two stronger lists. This leaves case D (1 + 3, 2), where the strongest and the weakest list unite against the median list.

In the 2008 Bavarian local elections there were 585 communities where just three lists campaigned. Of these, 513 fell into the apparentement-free category A, while fifty-one communities featured partition B (1, 2 + 3), six C (1 + 2, 3), and fifteen D (1 + 3, 2).

Table 3 shows the seat biases incurred by the partitions A–D when the divisor method with rounding down (D′Hondt) is used. The empirical values are the averages, among the communities where in 2008 the partition occurred, of the D′Hondt apportionment from the (unbiased) allocation of the divisor method with standard rounding (Webster/Sainte-Laguë). Most often the latter yields the same seat allocation as does the quota method with residual fit by largest remainders (Hamilton/Hare).

The theoretical values are the means calculated using the formulas in Table 2. Empirical and theoretical values conform quite satisfactorily. The total bias (=sum of all positive seat biases) is dampened in case B, as compared to the apparentement-free case A, enlarged in case C, and maximized in case D.

Practicalities defy theoretical predictions. In the 2008 Bavarian local elections it happened not once, but several times that the strongest list secured a double bonus by teaming up with weaker parties.

4 Large Parties Uniting with Small Parties

Table 4 presents an example of a double bonus, in Unterallgäu County. The divisor method with standard rounding (Webster/Sainte-Laguë) allocates about one seat per each 51,900 votes. The strongest list, with quotient 1,377,975/ 51,900 = 26.55, is allocated 27 seats (Column A). Even with no apparentement, the divisor method with rounding down (D′Hondt) gives an advantage by awarding it 28 seats since, with electoral key 48,000, the quotient 1,377,975/48,000 = 28.7 is rounded down (Column B).

However, two list apparentements had been registered. The strongest List 1 united with the fifth-strongest List 5, and the forth- and sixth-strongest lists joint together. Table 4 shows what happened. Without list apparentements, List 1 and 5 would have gained 28 + 3 = 31 D′Hondt seats. With list apparentements, they won 32 seats (Column C1). The sub-apportionment assigns the second bonus seat to the stronger of the two partners, List 1 (Column C2). In the end List 1 is apportioned 29 seats, rather than its unbiased share of 27 seats (Column D).

5 Lottery Effects

Formation of list apparentements turns into a lottery for the reason that there is a plethora of ways as to how a set of lists may be partitioned into different apparentements. The six lists in Table 2 admit 201 apparentements; for the seven lists of Table 4 the count[6] grows to 875. The information for voters that "some lists form an apparentement" is much too vague to be of any value. The abundance of possible list apparentements makes it impossible to intuitively assess their consequences.

A first rule applies to list apparentements just as it applies to any other game: *Nothing ventured, nothing gained.* Lists who prefer to maintain their independence and do not join an apparentement must, on average, endure a seat deficit so that their competitors may be served with a seat surplus.[7]

The second rule is a counterpart of the first: *If there is just one list apparentement, its partners are guaranteed to be on the winner's side*. On the average the partners of a sole apparentement receive a seat surplus as compared to the apparentement-free D′Hondt apportionment.[8] In 2008 two thirds of the Bavarian

---

[6] Our counts neglect the borderline cases (1) "everyone stands alone" (1, 2,…, ℓ−1, ℓ) and there is no sub-apportionment, and (2) "all join together" (1 + 2+… + "ℓ−1" + ℓ) and there is no super-apportionment.

[7] The formulas from Table 2 yield $\mathrm{D'H}(j \mid L_1, \ldots, L_k) - \mathrm{D''H}(j) = -(\ell - k)\, s(j)\,/\,2 < 0$, assuming that List j remains alone while other lists enter into an apparentement of two or more partners ($k < \ell$).

[8] The formulas give $\mathrm{D'H}(j \mid V;\ \{\, i \,\},\, i \notin V) - \mathrm{D''H}(j) = \ (1 - s(V))\,(p-1)\,s(j)\,/(2\, s(V)) > 0$, assuming List j is one of p partners of the (sole) list apparentement V, the other $\ell - p$ lists running by themselves.

Table 4 Double bonus for the strongest list, Unterallgäu County 2008

List (apparentement)	Votes	(A) S-L = H/H w/o a.	(B) D′H w/o a.	(C1) D′H with a.	(C2) D′H 1 + 5	(C3) D′H 4 + 6	(D) Final seats
List 1	1,377,975	27	28		29		29
List 2	730,846	14	15	14			14
List 3	337,937	7	7	6			6
List 4	189,648	4	3			4	4
List 5	181,235	3	3		3		3
List 6	163,465	3	3			3	3
List 7	85,511	2	1	1			1
Apparentement 1 + 5	(1,559,210)			32			
Apparentement 4 + 6	(353,113)			7			
Sum	3,066,617	60	60	60	32	7	60
Divisor		51,900	48,000	48,724	47,000	44,000	

communities (456 of 668, see Table 1) featured just one list apparentement. Its partners could look forward to a bonus simply because their competitors were napping.

In 212 communities, however, two or more list apparentements were registered. These are the instances when the elections turn into a lottery. Surpluses and deficits constitute a zero-sum game. It is plainly impossible that each and every protagonist finishes up with a bonus. But who is advantaged, and who is disadvantaged, is predictable only after extensive calculations, and in practice turns into mere luck.

It is not even recognizable what happens to the total bias of the system. It is a wide-spread belief that list apparentements *always* dampen the total bias. This belief is erroneous, as has already been seen in Table 3. Moreover, of the 201 apparentements into which six lists may be partitioned, 73 were realized in the 2008 Bavarian local elections. Of these, barely 44 diminished the total bias. With the other 29 partitions—that is, in more than a third of all cases—the total bias became larger, not smaller.

Here is a seemingly balanced example worth mulling over, from the previous Bavarian local elections in 2002. In Bad Füssing (AGS 09275116) nine lists campaigned, and formed three apparentements of three partners each, namely $1 + 3 + 5$, $2 + 4 + 7$ und $6 + 8 + 9$. Again lists are numbered according to their ranking by votes received. Who paid the bill? Who made the best cut? In case the gentle reader would like to ponder the example, we masquerade the answers as reference Xyz (2002).

6 Discordant Seat Assignments

We consider it a system defect when the popular vote is turned upside down, and fewer votes finish up with more seats. We call a setting in which of two lists that one with fewer votes gets more seats, a *discordant seat assignment*, or simply a discordance.

Table 5 further elaborates on the Friedberg example of Table 2, illustrating how discordances evolve. The second-strongest list ranks by more than five-thousand votes behind the winning list. Yet List 2 wins 13 seats, while List 1 acquires only 12. The theoretical formulas in Table 2 already foreshadowed this mishap.

Table 6 assembles all thirty-six discordances which emerged during the 2008 Bavarian local elections.[9] The Friedberg example is not a singular exception. In seven instances the second-strongest list leapt to the top as far as seats are concerned; while the strongest list dropped down to rank two. In Eurasburg (AGS

[9] Vote counts reflect council sizes, as every voter has as many votes as there are council seats to fill.

Table 5 Discordance victory of the second-strongest list, City of Friedberg 2008

List (apparentement)	Votes	(A) S-L w/o a.	(B) D′H w/o a.	(C) H/H w/o a.	(D1) D′H with a.	(D2) D′H 2 + 3+5	(E) Final seats
List 1	150,615	12	12	12	12		12
List 2	145,292	12	12	11		13	13
List 3	30,558	2	2	2		2	2
List 4	28,428	2	2	2	2		2
List 5	18,291	1	1	2		1	1
List 6	12,010	1	1	1	0		0
Apparentement 2 + 3+5	(194,141)				16		
Sum	385,194	30	30	30	30	16	30
Quota|Divisor		1,2000	1,2400	12,839.8	1,2100	1,1000	

Table 6 Discordant seat assignments, Bavaria 2008

AGS	Community|Council	Size	D′Hondt discordances	List partitionings
09173123	Eurasburg	16	(1)6206 − 3:(2)6172 − 4	2 + 6 + 7, 3 + 4, 1, 5, 8
09175122	Grafing	24	(3)20229 − 3:(4)19282 − 4	1, 2 + 4, 3, 5
09179121	Fürstenfeldbruck	40	(4)43625 − 3:(5)42345 − 4	2 + 3 + 5 + 6, 1, 4
09179145	Puchheim	24	(4)19148 − 2:(5)19146 − 3	1, 2, 3 + 5, 4
09180125	Oberammergau	20	(7)4498 − 1:(8)4321 − 2	1 + 3 + 7, 4 + 5 + 6 + 8 + 9, 2
09180129	Saulgrub	12	(1)5267 − 5:(2)5137 − 6	2 + 3, 1
09186	Pfaffenhofen County	60	(5)169961 − 3:(6)169250 − 4	1, 2, 3, 6 + 7, 4, 5
09186122	Geisenfeld	20	(5)4382 − 0:(6)4032 − 1	1 + 4, 2 + 6, 3, 5
09272118	Freyung	20	(3)6291 − 1:(5)5950 − 2 (4)5953 − 1:(5)5950 − 2	1 + 7, 2, 5 + 6, 3, 4
09273	Kelheim County	60	(5)162316 − 3:(6)152563 − 4	1 + 4 + 5, 2, 3, 6 + 7 + 8
09277111	Arnstorf	20	(2)6300 − 2:(3)5692 − 3	3 + 4 + 5 + 6 + 8, 1 + 7 + 9, 2
09279128	Moosthenning	16	(2)5574 − 2:(3)5201 − 3	3 + 5 + 7, 1, 2, 4, 8 + 9, 6
09376163	Schwarzenfeld	20	(2)8023 − 2:(3)7678−3	1 + 5, 3 + 4 + 6 + 7, 2
09376169	Stulln	12	(2)4642 − 2:(3)3962 − 3	1, 3 + 4, 2
09472143	Goldkronach	16	(3)4866 − 2:(4)4577 − 3	4 + 5 + 6, 1 + 8, 2, 3, 7
09472167	Mistelgau	16	(1)6922 − 3:(2)6016 − 4	2 + 3 + 4 + 5 + 6 + 7 + 8, 1
09472197	Waischenfeld	16	(4)2723 − 1:(5)2394 − 2	1, 5 + 6, 3 + 8, 2, 4, 7, 9
09474121	Ebermannstadt	20	(4)9525 − 2:(5)8022 − 3	1 + 7, 3 + 5 + 6, 2, 4
09474123	Eggolsheim	20	(2)8468 − 2:(3)8289 − 3	1 + 6, 3 + 4, 5 + 7 + 9 + 10, 2, 8
09474129	Gößweinstein	16	(2)7611 − 3:(3)7529 − 4	1 + 5, 3 + 4, 2
09572111	Adelsdorf	20	(2)21082 − 5:(3)20852 − 6	1 + 4, 3 + 5, 2
09673172	Sulzdorf	12	(4)2263 − 1:(5)1984 − 2	2 + 3 + 5, 1, 4, 6 + 7, 8
09673173	Sulzfeld	12	(4)3322 − 1:(5)3314 − 2	1, 3 + 5, 2, 4
09678170	Röthlein	16	(3)5885 − 2:(4)5704 − 3	1 + 6, 4 + 5, 2, 3, 7
09678193	Werneck	24	(3)13129 − 2:(4)11690 − 3	4 + 7 + 8 + 10 + 11, 2 + 6 + 9, 3 + 12 + 13, 1, 5

(continued)

Table 6 (continued)

AGS	Community|Council	Size	D′Hondt discordances	List partitionings
09771130	Friedberg	30	(1)150615 − 12:(2)145292 − 13	2 + 3 + 5, 1, 4, 6
09772147	Gersthofen	30	(2)42811 − 5:(3)41234 − 6	1, 3 + 4 + 5 + 6, 2
09772177	Meitingen	24	(2)18634 − 3:(3)17483 − 4	1 + 6, 3 + 4 + 5 + 7, 2
09772178	Mickhausen	12	(1)2416 − 4:(2)2288 − 5	2 + 3, 1
09773117	Bissingen	16	(4)3577 − 1:(5)3346 − 2	1 + 3 + 5 + 7 + 10, 2 + 6 + 8 + 9, 4
09774135	Günzburg	24	(1)47078 − 7:(2)44800 − 8	1 + 4 + 5, 2 + 3 + 6
09774171	Offingen	16	(1)10916 − 5:(2)10540 − 6	2 + 5, 1, 3, 4
09777129	Füssen	24	(2)24004 − 4:(3)22502 − 5	3 + 4 + 6 + 7 + 9 + 10, 1 + 5, 2 + 8
09779147	Fremdingen	14	(3)6748 − 2:(4)6597 − 3	2 + 4 + 5, 1, 3
09780117	Buchenberg	16	(3)3041 − 1:(4)2642 − 2	1, 2, 4 + 5, 3

09173123) the strongest list (1) with 6,206 votes got 3 seats, while the second-strongest list (2) fell back in votes (6,172), but jumped ahead in seats (4).

The partitioning of the apparentements are exhibited in the right-most column of Table 6, demonstrating the abundance of possibilities of who may go together with whomever else. Eurasburg featured two apparentements. The second-, sixth-, and seventh-strongest lists united (2 + 6 + 7), and finished first in terms of votes. The third- and fourth-strongest lists (3 + 4) came in second. The others stood alone (1, 5, 8).

Since list apparentements entail repeated apportionment calculations with multiple steps of rounding, *every* electoral formula is prone to discordant seat assignments. In particular, neither the divisor method with standard rounding (Webster/Sainte-Laguë) nor the quota method with residual fit by largest remainders (Hamilton/Hare) are immune to discordances. However, due to its notorious seat biases the D′Hondt method gives rise to discordances about twice as often as compared to its unbiased competitors. While D′Hondt systematically favors the stronger partners within an apparentement, the unbiased methods behave unpredictably and, when producing discordances, may favor lists within the apparentements, or lists that stand alone.

7 Constitutional Principles

May local elections turn into a lottery? Article 28 of the *Grundgesetz,* the German constitution, defines the standard. Elections in Germany must be universal and direct, as well as free, equal, and secret. The principle of electoral equality acquires a double meaning, *Chancengleichheit der Parteien* (equal chances for parties) aiming at parties and candidates, and *Erfolgswertgleichheit der Stimmen* (equal success values of votes) honoring the role of voters.

The lottery character of list apparentements certainly honors the equality principle as far as equal chances for parties are concerned. Officials of all parties have an equal opportunity to place their stakes in the game. If some players miss their turn, as in Friedberg Lists 1, 4 and 6, such negligence does not render the law unconstitutional.

We believe that constitutionality of list apparentements is much more problematic when considered from the voters' point of view. It is questionable whether the election can rightly claim to be direct. After all, two apportionment calculations are called for, and this detour hardly qualifies as a direct route from votes to seats.

Furthermore we find it more than unclear whether votes can be considered free. From the voters' viewpoint it is unknown third parties who interfere and decide whether the votes first undergo a preliminary evaluation via list apparentements, or not.

And what about electoral equality? If the constitution requires all votes to achieve an equal success value, how does it happen that fewer votes can lead to more seats? In order to justify such a contradiction, a sophistic vindication is called for that we are unable to offer with our modest talents as statisticians.[10]

References

Bochsler, D. (2009). Who gains from apparentements under D′Hondt? *Electoral Studies*, *29*(2010), 617–627.

Gallagher, M., & Mitchell, P. (Eds.) (2005). *The politics of electoral systems.* Oxford.

Leutgäb, P. (2008). Listenverbindungen, Thesis, Institut für Mathematik, Universität Augsburg.

Pukelsheim, F. (2002). Die Väter der Mandatszuteilungsverfahren. *Spektrum der Wissen-schaft, 9/2002*, 83.

Pukelsheim, F. (2003). Erfolgswertgleichheit der Wählerstimmen? Der schwierige Umgang mit einem hehren Ideal. *Stadtforschung und Statistik, 1/2003*, 56–61.

Pukelsheim, F., & Schuhmacher, C. (2004). Das neue Zürcher Zuteilungsverfahren für Parlamentswahlen. *Aktuelle Juristische Praxis—Pratique Juridique Actuelle 5/2004*, 505–522.

Pukelsheim, F., Maier, S., & Leutgäb, P. (2009). Zur Vollmandat-Sperrklausel im Kommunalwahlgesetz. *Nordrhein-Westfälische Verwaltungsblätter, 3/2009*, 85–90.

Schuster, K., Pukelsheim, F., Drton, M., & Draper, N. R. (2003). Seat biases of apportionment methods for proportional representation. *Electoral Studies, 22*(2003), 651–676.

Xyz (2002). List 1 paid a major portion of the bill, its bonus shrank by a third of a seat ($0.548 - 0.914 = -0.366$). List 6 benefitted most, gaining close to half a seat ($0.202 - (-0.227) = 0.429$). In the 2002 election, List 6 won two seats, and thus got ahead of List 5 who had to resign themselves to just one seat.

[10] The German Federal Constitutional Court shares a critical stand on list apparentements: *Every list apparentement* [*leads*] *to a violation of the principle of electoral equality, since votes are assigned unequal weights without justifying the deviation from equality by a forceful, substantive argument,* see BVerfGE 82 (1991) 322–352 [345]. The decision concerned the by-passing of the five-percent hurdle in the first all-German elections in 1990, not the role of list apparentements in local elections.

Voting and Power

Annick Laruelle and Federico Valenciano

1 Introduction

Since the seminal contributions of Penrose (1946), Shapley and Shubik (1953) and Banzhaf (1965), the question of the 'voting power' in voting situations has received attention from many researchers. Variations, different characterizations and alternative interpretations of the seminal concepts, their normative implications for the design of collective decision making procedures and innumerable applications can be found in the game theoretic and social choice literature.[1] Nevertheless, even half a century later and despite the proliferation of contributions in the field, one cannot speak of consensus in the scientific community about the soundness of the foundations of this body of knowledge and consequently about the normative value of the (often contradictory) practical recommendations

An earlier version of this chapter has been published in Essays in Honor of Hannu Nurmi (Homo Oeconomicus 26, 2009), ed. by M.J. Holler and M. Widgrén. This research is supported by the Spanish Ministerio de Ciencia e Innovación under project ECO2009-11213, co-funded by the ERDF, and by Basque Government funding to Grupo Consolidado GIC07/146-IT-377-07.

[1] We do not give a comprehensive or even a summary list of the contributions in the field as we believe it to be unnecessary for the purpose of this chapter.

A. Laruelle
BRiDGE, Foundation of Economic Analysis I, University of the Basque Country (UPV/EHU), Avenida Lehendakari Aguirre, 83 48015 Bilbao, Spain

A. Laruelle (✉)
IKERBASQUE Basque Foundation of Science, 48011 Bilbao, Spain
e-mail: a.laruelle@ikerbasque.org

F. Valenciano
BRiDGE, Applied Economics IV, University of the Basque Country (UPV/EHU), Avenida Lehendakari Aguirre, 83 48015 Bilbao, Spain
e-mail: federico.valenciano@ehu.es

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_8,

that stem from it. Because of this lack of consensus it would be exaggerating to speak of a voting power 'paradigm'[2] in the sense of Kuhn (1970).

Our purpose here is not to survey discrepancies between competing approaches within this field but to make some suggestions for providing more convincing foundations, still within the a priori point of view but widening the conceptual framework. The ideas summarized in this chapter are developed in detail in Laruelle and Valenciano (2008c).

The rest of the chapter is organized as follows. In Sect. 2, we give summarily our way of approaching the question and convey the main ideas. In the next two sections we concentrate on each of the two basic scenarios in which a committee may have to make decisions. In Sect. 3 we consider *'take-it-or-leave-it'* situations, and in Sect. 4 we consider *'bargaining committees'* situations. In Sect. 5 we summarize the main conclusions.

2 Suggestions for Clearer Foundations

The specification of a voting rule for making dichotomous choices (acceptance/rejection) neither involves nor requires the description of its users or 'players'. It suffices to specify the vote configurations that would mean acceptance and those that would mean rejection. The same voting rule can be used by different sets of users to make decisions about different types of issue. But describing a voting rule as a simple TU game[3] falls into the trap of producing a game where there are no players. When the (then) recently introduced Shapley (1953) value was applied to this game, the Shapley and Shubik (1954) index resulted. Shapley and Shubik interpret their index as an evaluation of a priori 'voting power' in the committee. As the marginal contribution of a player to a coalition in this game can only be 0 or 1, and is 1 only when the presence/absence of a player in a coalition makes it winning/losing, they also propose an interpretation in terms of likelihood of being pivotal or decisive. Hence the seminal duality or ambiguity:

Q.1: *Is the Shapley-Shubik index a 'value', that is, an expected payoff in a sort of bargaining situation, or an assessment of the likelihood of being decisive?*

Later Banzhaf (1965) takes the point of view of power as decisiveness, and criticizes the Shapley–Shubik index in view of the unnatural probability model underlying its probabilistic interpretation in the context of voting. If power means

---

[2] It may possibly be more accurate to use the term 'preparadigm' to refer to a set of concepts and ideas accepted and shared by a certain number of researchers within the academic world, related to the power indices literature, which is basically surveyed and systematized in Felsenthal and Machover's (1998). In fact, their book can be seen as the closest thing to an embodiment of that (pre)paradigm. By contrast, another group in the profession dismisses voting power literature altogether as irrelevant in view of its weak foundations.

[3] By assigning 'worth' 1 to all 'winning coalitions' and 0 to the others.

being decisive, then a measure of power is given by the probability of being so. Thus an a priori evaluation of power, if all vote configurations are equally probable a priori, is the probability of being decisive under this assumption. Note that Penrose in 1946 independently reached basically the same conclusion in a narrower formal setting. In fact Banzhaf only 'almost' says so, as he destroys this clean probabilistic interpretation by 'normalizing' the vector (i.e. dividing the vector of probabilities by its norm so as to make its coordinates add up to 1). So the old dispute is served up:

Q.2: *Which is better as a measure of voting power: the Shapley-Shubik index or the Banzhaf index? What are more relevant: axioms or probabilities?*

It is our view that in order to solve these dilemmas and dissipate ambiguities a more basic issue must be addressed: *What are we talking about?* Instead of starting with abstract terms such as 'power' related to an exceedingly broad class of situations (any collective body that makes decisions by vote) and getting entangled prematurely in big words, we think it wiser to

(i) *start by setting the analysis of voting situations as the central goal.*

Collective decision-making by vote may include an extremely wide, heterogeneous constellation of voting situations: law-making in a parliament, a parliament vote for the endorsement of a government, a referendum, a presidential election, decision-making in a governmental cabinet, a shareholders' meeting, an international, intergovernmental or other council, etc. By setting the analysis of voting situations as the central goal instead of the abstraction 'voting power', we mean to

(ii) *start from clear-cut models of well specified clear-cut voting situations,*

instead of starting from words denoting poorly-specified abstractions in poorly-specified situations. For instance, a committee capable of bargaining a proposal before voting is not the same as one only allowed to accept or reject proposals by vote. Millions of voters are not the same as a few, etc.

A dichotomy consistent with the above principles should distinguish between two types of voting situations or committees which make decisions under a voting rule: 'take-it-or-leave-it' committees and 'bargaining' committees.

(ii.1) A *'take-it-or-leave-it' committee* votes on different independent proposals over time, which are submitted to the committee by an external agency, so that the committee can only accept or reject proposals, but cannot modify them.

(ii.2) A *'bargaining' committee* deals with different issues over time, so that for each issue a different configuration of preferences emerges among its members over the set of feasible agreements, and the committee bargains about each issue in search of an agreement, to attain which it is entitled to adjust the proposal.

Though in reality it is often the case that a same committee acts sometimes as a 'take-it-or-leave-it' committee, and sometimes as a 'bargaining' committee, or even as something in between, this crisp differentiation of two clear-cut types of situation provides benchmarks for a better understanding of many mixed real world situations. As shown in the next two sections, they require different models and raise different questions with different answers which give rise to different recommendations. This neat differentiation and the analysis therein also sheds some light on the old ambiguities.

An ingredient common to both types of committee is the voting rule that governs decisions. In order to proceed a minimum of notation is needed. If n voters, labelled by $N = \{1, 2, \ldots, n\}$ are asked to vote 'yes' or 'no' on an issue, any result of a vote, or *vote configuration,* can be summarized by the subset of voters who vote 'yes': $S \subseteq N$. An *N-voting rule* is then specified by a set $W \subseteq 2^N$ of *winning* vote configurations such that (i) $N \in W$; (ii) $\emptyset \notin W$; (iii) If $S \in W$, then $T \in W$ for any T containing S; and (iv) If $S \in W$ then $N \backslash S \notin W$.

3 'Take-it-or-Leave-it' Committees

3.1 The 'Take-it-or-Leave-it' Environment

A pure 'take-it-or-leave-it' environment is that of a committee that makes dichotomous decisions (acceptance/rejection) by vote under the following conditions: (i) the committee votes on different independent proposals over time; (ii) proposals are submitted to the committee by an external agency; (iii) the committee can only accept or reject each proposal, but cannot modify them; and (iv) a proposal is accepted if a winning vote configuration according to the specifications of the voting rule emerges.

In real world committees these conditions are seldom all satisfied. Nevertheless, for a sound analysis it is necessary to make explicit and precise assumptions about the environment, and this is the only way to have clear conclusions. Indeed, traditional power indices and the credibility of voting power theory is undermined by the lack of clarity about the precise specification of the underlying collective decision-making situation.

Under the above conditions, it seems clear that, except in the case of indifference about the outcome, each voter's vote is determined by his/her preferences. In particular these conditions leave no margin for bargaining. The impossibility of modifying proposals, their independence over time, etc., rule out the possibility of bargaining and consequently of strategic behavior. In other words, decision-making in a take-it-or-leave-it committee is *not* a game situation, therefore *in a pure 'take-it-or-leave-it' environment game-theoretic considerations are out of place.*

3.2 The Basic Issue in a 'Take-it-or-it-Leave-it' Environment

As briefly commented in the introduction, there are certain ambiguities at the very foundations of traditional voting power theory, concerning the precise conditions under which the collective decision-making process takes place, and concerning the interpretation of some power indices. Either explicitly or implicitly the notion of power as decisiveness, i.e. the likelihood of one's vote being in a position to decide the outcome, underlies most traditional voting power literature.

Were it not for the weight of theoretical inertia it would hardly be necessary to argue about the irrelevance of this notion in a pure 'take-it-or-leave-it' environment, where voting behavior immediately follows preferences.[4] *Decisiveness can be a form or, more precisely, a source of power only in a situation in which there is room for negotiation and the possibility of using it with this purpose. But the conditions that specify a 'take-it-or-leave-it' environment preclude that possibility.* For instance, voting on an issue against one's preferences in exchange for someone else doing the same in one's favor on a different issue is not possible in case of strict independence between proposals, as assumed.

The interest of any voter lies in obtaining the desired outcome, and in a 'take-it-or-leave-it' situation nothing better can be done in order to achieve that end than just voting accordingly. Thus, *having success or satisfaction (i.e. winning the vote) is the central issue in this kind of voting situation.* If so the assessment of a voting situation of this type with normative purposes requires us to assess the likelihood of each voter having his/her way. For an a priori assessment it seems natural to assume all configurations of preferences or equivalently (at least if no indifference occurs) all vote configurations as being equally probable. This assumption underlies a variety of power indices in the literature, that is, the probability of every vote configuration S is $\frac{1}{2^n}$.

Assuming this probabilistic model, the probability of a voter i being successful in a vote under a voting rule W is given by

$$Succ_i(W) := \text{ Prob}(i \text{ is successful}) = \sum_{S: i \in S \in W} \frac{1}{2^n} + \sum_{S: i \notin S \notin W} \frac{1}{2^n}. \tag{1}$$

As commented above, more attention has been paid to the probability of being decisive under the same probabilistic model, given by

$$Dec_i(W) := \text{ Prob}(i \text{ is decisive}) = \sum_{\substack{S \,:\, i \in S \in W \\ S \backslash i \notin W}} \frac{1}{2^n} + \sum_{\substack{S \,:\, i \notin S \notin W \\ S \cup i \in W}} \frac{1}{2^n}. \tag{2}$$

---

[4] Rae (1969), Brams and Lake (1978), Barry (1980) (from whom we take the term 'success'), Straffin, Davis, and Brams (1982), and more recently König and Bräuninger (1998) all pay attention to the notion of success or satisfaction, but the ambiguity has remained unsolved due to the lack of definite clarification about the underlying voting situation. In Laruelle, Martínez and Valenciano (2006) we argue in support of success as the relevant notion in a 'take-it-or-leave-it' committee.

Other conditional variants can be considered.[5] See Laruelle and Valenciano (2005).

3.3 Normative Recommendations

If the relevant issue in a 'take-it-or-leave-it' environment is the likelihood of success, then that should be the basis for normative recommendations. The question arises of what recommendations can be made for the choice of the voting rule in a committee of this type in which each member acts on behalf of a group of a different size. There are two basic points of view for the basis of such recommendations: egalitarian (equalizing the expected utility of all individuals represented) and utilitarian (maximizing the aggregated expected utility of the individuals represented). The implementation of either principle with respect to the people represented requires some assumption about the influence of those individuals on the decisions of the committee, *and* about the voters' utilities at stake. The well-known idealized two-stage decision process assuming that each representative follows the majority opinion of his/her group on every issue can be neatly modeled by a composite rule in which decisions are made directly by the people represented. As to the voters' utilities, a symmetry or anonymity assumption seems the most natural for a normative approach. This approach is taken in Laruelle and Valenciano (2008c), and the conclusions are the following.

3.3.1 Egalitarianism and the Square Root Rule

The egalitarian principle, according to which all individuals should have an equal expected utility, would be implemented[6] by a voting rule in the committee for which the Banzhaf index of each representative is proportional to the square root of the size of the group that s/he represents.

But this is the well-known 'square root rule' (SQRR) that appears in voting power literature,[7] where power is understood as the probability of being decisive, as a means of equalizing 'voting power'. So, are we back to the old recommendation? Yes and no, but mainly no. The important difference is the following: in the voting power approach, where the likelihood of being decisive is what matters, the SQRR is the way of implementing the egalitarian principle *in terms of power*

[5] Actually some of them have been considered in the literature, but those that we consider most relevant have so far been overlooked, namely, the probabilities of a voter being successful *conditional* either upon his/her voting 'yes' or upon his/her voting 'not'. They are calculated in Laruelle, Martínez, and Valenciano (2006) for some voting rules in the Council of Ministers of the EU with interesting results.

[6] With close approximation if all the groups are large enough.

[7] Conjectured by Morriss (1987) and rigorously proved by Felsenthal and Machover (1999).

so understood. But in the utility-based approach it is merely a sufficient condition for (approximately) equalizing the expected utility of all the individuals represented. Usually the SQRR can only be implemented approximately, so exact fulfillment is exceptional. Herein lies the crucial difference between the two approaches: according to the traditional approach the differences between the Banzhaf indices of individuals from different groups are seen as differences in the substantive notion to be equalized (i.e. power as decisiveness), while in the utility-based approach the substantive differences are in utilities. It turns out that when groups are big enough the *expected utilities of all individuals are very close, whatever the voting rule in the committee*.

3.3.2 Utilitarianism and the 2nd Square Root Rule

The utilitarian principle, according to which the sum of all individuals expected utilities should be maximized, would be implemented by a weighted voting rule that assigns to each representative a weight proportional to the square root of the group sizes, and the quota is half the sum of the weights.[8]

Again we are back to a well-known recommendation: the ‘second square root rule’ (2nd SQRR). Nevertheless it is worth remarking that the underlying justifications are different and the differences are important. As in the case of the recommendation based on the egalitarian point of view, the utilitarian recommendation has a clear justification only for this special type of committee, while for the more complex case of bargaining committees both recommendations lack a clear basis. The lack of a precise specification of the voting situation in the traditional analysis has so far concealed this important point. As we will see in a bargaining committee these recommendations lack foundations and this has important consequences for applications.

4 Bargaining Committees

4.1 The Bargaining Environment

As soon as any of the conditions specifying what has been called a ‘take-it-or-leave-it’ environment is relaxed the situation changes drastically. For instance, if decisions on different issues cease to be independent, bargaining over votes of the form ‘I will vote on this issue against my preferences in exchange for your doing

[8] With close approximation if all the groups are large enough, and it is assumed that all vote configurations are equally probable and that the utility of winning a vote is the same in case of acceptance as in the case of rejection (if voters place different values on having the desired result when they support approval and when they support rejection, the quota should be adjusted).

the same in my favor on that issue' becomes possible. Also, if the committee can modify proposals negotiation prior to voting is to be expected. In fact, outside the rather constrained environment specified as 'take-it-or-leave-it' there are many possibilities.

Thus the '*non*-take-it-or-leave-it' scenario is susceptible to a variety of specifications that can be seen as variations of the 'bargaining' environment. The analysis is possible only if the conditions under which negotiation takes place are specified. To fix ideas, although others are possible, here we discuss the following specification of a bargaining committee that makes decisions using a voting rule under the following conditions: (i) the committee deals with different issues over time; (ii) for each issue a different configuration of preferences emerges among the members of the committee over the set of feasible agreements concerning the issue at stake; (iii) the committee bargains about each issue in search of a consensus, to which end it is entitled to adjust the proposal; and (iv) any winning coalition[9] can enforce any agreement.

Now the situation is much more complicated. The environment permits bargaining among the members of the committee, and consequently *behavior no longer trivially follows preferences. Now game-theoretic considerations are in order because the situation is inherently game-theoretic.*

4.2 The Basic Issues in a Bargaining Environment

First note that *in a bargaining situation the basic issue is that of the outcome* of negotiations. That is, given a preference profile of the members of the committee and a bargaining environment, what will the outcome be? Or at least, what outcome can reasonably be expected? It should be remarked that only if this basic question is answered can other relevant issues be addressed, e.g. the question of the influence of the voting rule on the outcome of negotiations and the question of the 'power' that the voting rule gives to each member. In particular, the meaning of the term 'power' in this context can only become clear when one has an answer to the first basic question.

In order to provide an answer to the central question a formal model of a bargaining committee as specified is needed. A model of such a bargaining committee should incorporate at least the following information: the voting rule under which negotiation takes place, and the preferences of the *players* (the usual game-theoretic term is now appropriate). Other elements need to be included for a more realistic model, but it is best to start with as simple a model as possible to see what conclusions can be drawn from it. In Laruelle and Valenciano (2007) a model of an n-person bargaining committee incorporating these two ingredients is introduced. The

[9] Notice that, unlike what happens in the case of a take-it-or-leave-it situation, in this context the old traditional game-flavored term "winning coalition" is appropriate.

first element is just the n-person voting rule W. As to the second, the players' preferences, under the same assumptions as in Nash (1950), i.e. assuming that they are expected utility or von Neumann-Morgenstern (1944) preferences, they can be represented à la Nash by a pair $B = (D, d)$, where D is the set[10] of feasible utility vectors or 'payoffs', and d is the vector of utilities in case of disagreement.

In this two-ingredient setting the question of rational expectations about the outcome of negotiations can be addressed from two different game-theoretic points of view: the cooperative and the noncooperative approaches. The cooperative method consists of ignoring details concerning the way in which negotiations take place, and 'guessing' the outcome of negotiations between ideally rational players by assuming reasonable properties of the map that maps 'problems' (B, W) into payoffs $\Phi(B, W) \in R^N$. The most influential paradigm of the cooperative approach is Nash's (1950) bargaining solution. In Laruelle and Valenciano (2007) Nash's classical approach is extended or adapted to this two-ingredient setting, *assuming that players in a bargaining committee bargain in search of unanimous agreement.* In this way, by assuming adequate adaptation to our setting of some reasonable conditions to expect for a bargaining outcome (efficiency, anonymity, independence of irrelevant alternatives, invariance w.r.t. affine transformations, and null player), it is proved that a general 'solution' (i.e. an N-vector valued map $(B, W) \longmapsto \Phi(B, W)$) should take the form

$$\Phi(B, W) = Nash^{\varphi(W)}(B) = \arg\max_{x \in D_d} \prod_{i \in N} (x_i - d_i)^{\varphi_i(W)}. \tag{3}$$

That is to say: a reasonable outcome of negotiations is given by the weighted Nash bargaining solution (Kalai 1977) where the weights are a function $\varphi(W)$ of the voting rule. Moreover this function must satisfy anonymity and null-player. Note also that these two properties are the most compelling ones concerning 'power indices'.[11] *Thus formula* (3) *sets the 'contest' between power indices candidates to replace* $\varphi(W)$ *in* (3) *in a new setting and provides a new interpretation of them in terms of 'bargaining power' in the precise game theoretic sense.*[12] In the same chapter it is shown how adding an adaptation of Dubey's (1975) lattice property to the other conditions singles out the Shapley–Shubik index in (3), i.e. in that case the solution is

$$\Phi(B, W) = Nash^{Sh(W)}(B) = \arg\max_{x \in D_d} \prod_{i \in N} (x_i - d_i)^{Sh_i(W)}. \tag{4}$$

---

[10] D is a *closed, convex* and *comprehensive* (i.e., $x \le y \in D \Rightarrow x \in D$) set in R^N containing d, such that there exist $x \in D$ s.t. $x > d$, and such that the set $D_d := \{x \in D : x \ge d\}$ is *bounded*.

[11] These properties are satisfied by the two most popular power indices, but also by all semivalues (Dubey, Neyman and Weber (1981), see also Laruelle and Valenciano (2001, 2002, 2003)).

[12] In game-theoretic terms, the weight associated with each player for an asymmetric Nash bargaining solution is called 'bargaining power', as reflecting the relative advantage or disadvantage that the environment gives to each player.

It is also interesting to remark that, as shown in Laruelle and Valenciano (2007), when the bargaining element, i.e. the preference profile in the committee summarized by $B = (D, d)$, is *transferable utility*-like, that is,

$$B = \Lambda := (\Delta, 0), \quad \text{where } \Delta := \left\{ x \in R^N : \sum_{i \in N} x_i \leq 1 \right\}$$

then we have that (3) and (4) become respectively:

$$\Phi(\Lambda, W) = Nash^{\varphi(W)}(\Lambda) = \bar{\varphi}(W) \tag{5}$$

$$\Phi(\Lambda, W) = Nash^{Sh(W)}(\Lambda) = Sh(W), \tag{6}$$

where $\bar{\varphi}(W)$ denotes the normalization of vector $\varphi(W)$, which as commented in the conclusions solves the power/payoff dilemma mentioned in Sect. 2.

As mentioned above there is also the noncooperative approach, in which the model should specify with some detail the way in which negotiations take place. This is neither simple nor obvious in a situation of which the only ingredients so far are the voting rule that specifies what sets of members of the committee have the capacity to enforce any agreement, and the voters' preferences. A noncooperative modeling must necessarily choose a 'protocol' to reach any conclusion. The question arises whether (3) or (4) have a noncooperative foundation. In Laruelle and Valenciano (2008a) this problem is addressed and noncooperative foundations to (3) and (4) are provided. Assuming complete information, a family of noncooperative bargaining protocols is modeled based on the voting rule that provides noncooperative foundations for (3), which appear in this light as limit cases. The results based on the noncooperative model evidence the impact of the details of protocol on the outcome, and explain the lack of definite arguments (i.e. axioms compelling beyond argument) to go further than (3). Nevertheless (4) emerges associated with a very simple protocol also based on the voting rule under which negotiations take place, thus providing some sort of focal appeal as a reference term for the Shapley–Shubik index as an a priori measure of bargaining power.

4.3 Normative Recommendations

The model summarized in the previous section can be taken as a base for addressing the normative question of the most adequate voting rule in a committee of representatives. Namely, if a voting rule is to be chosen for a committee that is going to make decisions in a bargaining environment as described: what rule is the most appropriate if each member acts on behalf of a different sized group?

Laruelle and Valenciano (2008b) addresses this problem, which is tricky because for each issue a different configuration of preferences emerges in the population represented by the members of the committee. Thus if by ‘appropriate’ we mean fair in some sense, nothing can be said unless some form or other of relation about the preferences within each group is assumed. By ‘fair’ we mean *neutral* in the following sense: a neutral voting rule for the committee is one such that all those represented are indifferent between bargaining directly (ideal and unfeasible, but theoretically tractable according to Nash’s classical bargaining solution[13]) and leaving it in the hands of a committee of representatives. This is obviously utopian, but it can be proved to be implementable under some ideal symmetry conditions for the preferences within each group. In real world situations this condition may well fail to occur in most cases, but this idealization seems a reasonable term of reference if a voting rule is to be chosen. In fact if certain conditions of symmetry within each group are assumed the following recommendation arises:

A neutral voting rule (Laruelle and Valenciano 2008b): *A neutral voting rule in a bargaining committee of representatives is one that gives each member a bargaining power proportional to the size of the group that he/she represents.*

Note that this recommendation is based on (3), i.e. it does not presuppose which is the right φ in formula (3), but notice all the same the difference from the square root rule recommendation. The neutral voting rule would be one for which:

$$\frac{\varphi_i(W)}{m_i} = \frac{\varphi_j(W)}{m_j} \qquad (\forall i, j \in N),$$

while according to the square root rule the fair voting rule is one for which

$$\frac{Bz_i(W)}{\sqrt{m_i}} = \frac{Bz_j(W)}{\sqrt{mj}} \qquad (\forall i, j \in N),$$

where Bz_i denotes voter i’s Banzhaf index.

5 Conclusions

Thus we have several conclusions. Consider Q.1 and Q.2 raised in Sect. 2. The mere statements of Q.1 and Q.2 now look confusing in themselves. The reason is the narrowness of the framework in which they were formulated.

First, in the light of the conceptual and formal framework summarized above it can be said that in a ‘take-it-or-leave-it’ committee the Banzhaf index seems an appropriate measure of a priori decisiveness founded on probabilistic terms,

[13] Note also that the Nash bargaining solution is a compromise between egalitarianism and utilitarianism.

although in such committees the relevant notion is that of success. Thus, in the light of the above analysis the popular square root rule, which enjoys ample support in view of its providing a priori equal chances of being decisive, appears ill-founded in two ways: first, it should be the goal of equalizing the likelihood of success that justifies it, but for this purpose the rule hardly matters; and, second and more importantly, *it is only in the context of 'take-it-or-leave-it' committees that this recommendation makes sense.* Most often supporters of this choice apply it to real world committees that make decisions in an environment closer to that of a bargaining committee than to that of a 'take-it-or-leave-it' committee.

Moreover, in the case of bargaining environments as specified here the Shapley–Shubik index, and indeed any other power index, when seen through the lens of formulae (4) or (3), appear as *the 'bargaining power' in the precise game-theoretic sense* (i.e. the weights of an asymmetric Nash bargaining solution) *that the voting rule gives to each member of the committee. Note that this bargaining power is related to decisiveness, but when the preference profile is TU-like the bargaining power of each player coincides with his/her expected payoff as given by* (5) and (6). Note that this solves the dilemma of Q.1. Among all those indices or measures φ which fit formula (3) the Shapley–Shubik index appears as a remarkable candidate for measuring the a priori bargaining power in such committees. Note also that in this case the support is not probabilistic but either cooperative-axiomatic or noncooperative game-theoretic (as a limit case).

Finally, there is the question of the quite different recommendations for the choice of voting rule in a committee of representatives depending on whether it is a 'take-it-or-leave-it' or bargaining committee. The different recommendations obtained from the analysis summarily described above seem rather disturbing, especially considering that real-world committees often act in an intermediate environment between the two *pure* types considered here. This does not invalidate these recommendations: on the contrary, a clear, sound conceptual founding only sets clear limits on the validity of the conclusions that one may get from formal models, while unclear situations underlying models and conceptual vagueness at the base of theory definitely blur the sense and validity of any conclusion.

References

Banzhaf, J. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review, 19*, 317–343.

Barry, B. (1980). Is it better to be powerful or lucky? Part I and Part II. *Political Studies, 28*(183–194), 338–352.

Brams, S. J., & Lake, M. (1978). Power and satisfaction in a representative democracy. In P. Ordershook (Ed.), *Game theory and political science* (pp. 529–562). New York: NYU Press.

Dubey, P. (1975). On the uniqueness of the Shapley value. *International Journal of Game Theory, 4*, 131–139.

Dubey, P., Neyman, A., & Weber, R. J. (1981). Value theory without efficiency. *Mathematics of Operations Research, 6*, 122–128.

Felsenthal, D. S., & Machover, M. (1998). *The measurement of voting power: Theory and practice, problems and paradoxes*. London: Edward Elgar Publishing.
Felsenthal, D. S., & Machover, M. (1999). Minimizing the mean majority deficit: The second square-root rule. *Mathematical Social Sciences*, *37*, 25–37.
Kalai, E. (1977). Nonsymmetric Nash solutions and replications of 2-person bargaining. *International Journal of Game Theory, 6*, 129–133.
König, T., & Bräuninger, T. (1998). The inclusiveness of European decision rules. *Journal of Theoretical Politics, 10*, 125–142.
Kuhn, T. S. (1970). *The structure of scientific revolutions, second edition enlarged*. Chicago: The University of Chicago Press. [First edition by The University of Chicago Press within the International Encyclopedia of Unified Science, Ed. by O. Neurath in 1962]
Laruelle, A., Martínez, R., & Valenciano, F. (2006). Success versus decisiveness: Conceptual discussion and case study. *Journal of Theoretical Politics*, *18*, 185–205.
Laruelle, A., & Valenciano, F. (2001). Shapley–Shubik and Banzhaf indices revisited. *Mathematics of Operations Research, 26*, 89–104.
Laruelle, A., & Valenciano, F. (2002). Power indices and the veil of ignorance. *International Journal of Game Theory, 31*, 331–339.
Laruelle, A., & Valenciano, F. (2003). Semivalues and voting power. *International Game Theory Review, 5*, 41–61.
Laruelle, A., & Valenciano, F. (2005). Assessing success and decisiveness in voting situations. *Social Choice and Welfare, 24*, 171–197.
Laruelle, A., & Valenciano, F. (2007). Bargaining in committees as an extension of Nash's bargaining theory. *Journal of Economic Theory, 132*, 291–305.
Laruelle, A., & Valenciano, F. (2008a). Non cooperative foundations of bargaining power in committees and the Shapley–Shubik index. *Games and Economic Behavior, 63*, 341–353.
Laruelle, A., & Valenciano, F. (2008b). Bargaining in committees of representatives: The 'neutral' voting rule. *Journal of Theoretical Politics, 20*, 93–106.
Laruelle, A., & Valenciano, F. (2008c). *Voting and collective decision-making: Bargaining and power*. Cambridge: Cambridge University Press.
Morriss, P. (1987). *Power: A philosophical analysis*. Manchester: Manchester University Press.
Nash, J. F. (1950). The bargaining problem. *Econometrica, 18*, 155–162.
Penrose, L. S. (1946). The elementary statistics of majority voting. *Journal of the Royal Statistical Society, 109*, 53–57.
Rae, D. (1969). Decision rules and individual values in constitutional choice. *American Political Science Review, 63*, 40–56.
Shapley, L. S. (1953). A value for n-person games. *Annals of Mathematical Studies, 28*, 307–317.
Shapley, L. S., & Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review, 48*, 787–792.
Straffin, P. D., Davis, M. D., & Brams, S. J. (1982). Power and satisfaction in an ideologically divided voting body. In M. J. Holler (Ed.), *Power, voting, and voting power* 239–253. Würzburg, Wien: Physica-Verlag.
von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton: Princeton University Press.

Decisiveness and Inclusiveness: Two Aspects of the Intergovernmental Choice of European Voting Rules

Thomas König and Thomas Bräuninger

1 The Constitutional Change of European Voting Rules

Constitutional events have recently changed the voting rules of the European (EU) legislature. Since the mid 1980s, Treaty reforms such as the Single European Act in 1987, the Maastricht Treaty in 1993 and the accession of Portugal and Spain (1986) and of Austria, Finland and Sweden (1995) have brought about continued modification to EU voting rules. To study this constitutional change two prominent measurement concepts are applied: the cooperative intergovernmental power index and the non-cooperative spatial model approach. Under the cooperative assumption of binding and enforceable agreements, the power index approach concentrates on different voting rules and the effects of accession scenarios. It reduces the phenomenon of Treaty reform to the question of how the distribution of voting weights in the Council of Ministers determines the distribution of power between member states.

The non-cooperative spatial model approach studies the strategic interaction between the Commission, the Council and—in some cases—the European Parliament (EP). Based on actors' spatial preferences, these models focus on the choice within rules in a uni- or multi-dimensional policy space. With regard to EU decision making their application reveals the strategic interaction between different voting bodies. Except for Article 148, 2b, all EU legislative procedures require a Commission proposal that must be adopted by the member states with unanimity, simple or qualified majority. Since most voting power studies have ignored the interaction between EU voting bodies, the spatial model approach calls

This chapter has been published in Homo Oeconomicus 17(1/2).

T. König (✉) · T. Bräuninger
University of Mannheim, 68131 Mannheim, Germany
e-mail: koenig@uni-mannheim.de

T. Bräuninger
e-mail: thomas.braeuninger@uni-mannheim.de

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_9,

their utility fundamentally into question (Garrett and Tsebelis 1996, p. 270). The application of spatial models, however, underestimates certain formal voting differences between member states, voting weights for example, when studying the impact of various EU decision-making procedures instead of the impact of member states' constitutional choice.

In this chapter we argue that both approaches have so far failed to give a satisfactory account of the complexity of the EU institutional framework. When member states make a constitutional choice, they decide on the application of voting rules for EU legislation without knowing their own spatial preferences on future legislative proposals (Buchanan and Tullock 1962, p. 78). This is the major difference between the choice within and the choice of voting rules. Intergovernmental power index analyses assume that the configuration of member states' voting weights in the Council of Ministers sufficiently explains the constitutional choice of voting rules in the expanding community (Brams and Affuso 1985; Hosli 1993; Johnston 1995; Widgrén 1994; Lane et al. 1995). If some actors are privileged with higher voting weights or individual veto rights, relative voting power studies calculate their formal prerogatives by their relative abilities of being decisive in forming winning coalitions. We call this element of constitutional choice analysis *relative decisiveness*, describing one property of voting rules, namely the distribution of expected gains from future decision making.

Yet, relative decisiveness does not reveal the second property of voting rules. According to this concept, all member states have the same relative power under simple majority, qualified majority and unanimous voting in the case of "One-Man-One-Vote" provisions (König and Bräuninger 1998, p. 136). Though weak simple majority voting increases the power to act of the voting body as a whole, unanimity requires the inclusion of all actors, thus leading to a high status quo bias (Buchanan and Tullock 1962; Coleman 1971). As the introduction of majority rules entails the possible exclusion of an actor, the crucial question is whether a member state accepts the possibility of being in a minority position in future EU legislation. In order to measure this second property of voting rules we introduce our concept of *absolute inclusiveness* describing the amount of expected gains from future decision making.

Our concern is the description of both aspects of power for the various EU procedures. In our view, member states choose specific voting rules to allocate power in order to obtain a (fair) distribution of legislative gains. We argue that the combination of both relative decisiveness and absolute inclusiveness reveals the allocation of power for specific policy areas. Relative decisiveness is understood as reflecting the actors' chances of determining the legislative outcome. Member states provide themselves with shares of votes to make a distribution of legislative gains they have agreed upon possible. Inclusiveness, however, refers to the possible number of decisions dependent on the strength of the voting rule. Both concepts are related to the member states' expectations of EU legislative gains, determining their constitutional choice of either unanimity, qualified majority, simple majority, or single veto players for specific policy areas.

In addition to previous voting power studies we not only take into account the power to act of the voting body as a whole but also the interaction between EU voting bodies. For this purpose we formulate *inter-institutional sets of winning coalitions* consisting of the Commission, members of the Council and—sometimes—of the EP. The remainder of this article is divided into three sections. In Sect. 2 we present our concept of acting entities and the inter-institutional set of winning coalitions. Thereafter, we introduce the indices on relative decisiveness and absolute inclusiveness. Finally, we apply both measures on current EU legislative sets of winning coalitions.

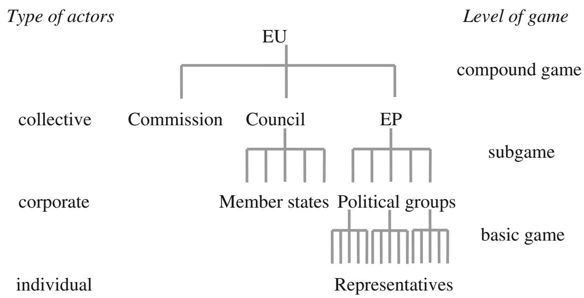
2 Acting Entities and Inter-Institutional Sets of Winning Coalitions

The concept of legislative winning coalitions is a fundamental element of the game-theoretical measurement of legislative entities' decisiveness and inclusiveness. Referring to the assumption on methodological individualism, both measures presuppose the identification of relevant actors and their procedural interaction. In the past, the identification problem of EU winning coalitions was often trivialised. This trivialisation found its expression in the ignoring of inter-institutional interaction, the assumption of a unitary (parliamentary) actor, or the disregard of actors' voting weights. We intend to improve the reliability of our approach by presenting our concept of EU actors and EU procedural interaction (Fig. 1).

In game-theoretical analyses, actors are simply defined as entities making choices in a specific context. This definition first presumes the identification of the acting entities and then considers the qualification of goal specificity, independence and consistence for their actions. In international relations theory, the unitary actor assumption of state behaviour is an illustrative example for the ongoing debate on the identification problem of acting entities (Achen 1995). In the field of power index analysis the primary task of actually identifying the relevant legislative entities is a well-known problem. "Paradoxes" like the paradoxes of quarrelling members, of new members and of size (Brams 1975), or the paradox of redistribution (Fischer and Schotter 1978) illustrate some of the crucial effects on relative decisiveness when either the set of entities, or the entities themselves, are modified.

To avoid identification problems, we begin our analysis by distinguishing between three types of legislative entities: individuals (natural persons), corporate actors (organisations with delegates as their agents), and collective actors (voting bodies). Like a natural person, a corporate actor is often considered to be a unitary entity having well-behaved preferences over outcomes and acting on purpose. Hence, there is no difference between individual and corporate actors if we ignore the controlling problem of delegates. In contrast to individual and corporate actors, collective actors are analysed as aggregates of individuals and/or corporate actors. The aggregation problem of individual and/or corporate actors is the topic of social

Fig. 1 EU legislative game



choice theory. Studies in this area show that the unitary actor assumption on collective actors rarely applies in cases of two or more preference dimensions (McKelvey 1979; Koehler 1990).

In EU legislation all three types of actors are relevant. Commission, Council and EP are voting bodies aggregating different sets of legislative entities. The *Commission* prepares proposals on which most of EU legislative decisions are based. In principle, the Commission is a college of twenty Commissioners each responsible for his or her General Directorate. Each Commissioner is provided with his or her own portfolio, carries the main leadership responsibility, and is independent of the Commission President in determining how to act on EU legislative decisions. We therefore conceptualise the Commission as a unitary actor in EU legislation with the responsible Commissioner as its agent (see also Spence 1994; Westlake 1994).

In the *Council*, the governments of the member states are represented by delegates mediating between their own governments and those of other delegates (Johnston 1994). National governments instruct their delegates, who then cast their votes homo-geneously in the Council. Since we ignore the controlling problem of delegation, we conceptualise the national delegate as an entity voting for its member state. Regarding the member states' votes, we can distinguish between equal and unequal settings. In the case of the EU qualified majority rule with 71.2 % voting quota, voting weights differ between large and smaller member states, thus providing for unequal settings. Against this, equal settings are provided for by the simple majority criterion and unanimity where member states are counted one-country-one-vote. Member states' votes are then aggregated in the Council, a collective actor facing other voting bodies in the course of EU legislative decision making.

Although the EP is less involved in EU legislative decision making, the disaggregation of the EP's entities causes further conceptual difficulties. Apart from different combinations of formal institutional settings, parliamentary systems differ in terms of specific peculiarities characteristic of a particular legislature. A specific characteristic of the EP is the affiliation of parliamentary representatives to both political and national groups. The fact that the vote of EP representatives on national group affiliation is merely a repetition of the intergovernmental, state-versus-state conflict in the Council, means that it is the political group affiliation that points out the unique contribution of parliamentary participation in EU legislation. We model political groups as EP entities with votes weighted according to their party representatives on the grounds that party cleavage is observed to dominate over national cleavage in the formation of majority coalitions (Jacobs et al. 1992; Attina 1990). Since no political group has an absolute majority at its disposal, political group votes are, by empirical necessity, aggregated in the EP.

The varying voting rules in the Council and the EP reveal different levels of EU legislation. We can distinguish between three levels: the basic game, the subgame, and the compound game. The basic game refers to the prime entities such as individuals or national party delegations which form the political groups in the EP. On the subgame level, the internal coalition problem of the member states in the Council and the parliamentary political groups in the EP has to be solved. Except for constitutional unanimity, the Council subgame offers two voting criteria, since, even in the case of majority voting, amendments always require unanimity among member states. Under the cooperation and codecision procedures, the EP may take action or no action. Preventing endorsement by no action slightly decreases the majority criterion, since—as the EP has always been a voting body consisting of an equal number of representatives—50 % of all votes are sufficient to prevent action while taking action needs 50 % plus one vote
Finally, the procedural settings of EU legislation define legislative sets of winning coalitions consisting of all entities necessary to adopt a proposal.[1] However, identifying EU *inter-institutional sets of winning coalitions* is made rather more complicated for two reasons: First, the Council and the EP's voting rules vary, and second, the role of the Commission is rather speculative. According to Article 155, the Commission holds the exclusive right to initiate legislation and the right to modify a proposal at any point of procedure (Article 189a, 2), thereby making the Commission the agenda setter. Moreover, the Commission also has the right to withdraw, if the proposal's original object is felt to have been emasculated by amendments (Usher 1994). The Commission cannot, therefore, be excluded from the set of all relevant legislative entities.

[1] EU legislative sets of winning coalitions require consent among all relevant voting bodies and thus depend on the solution of the coalition problems at the subgame level. Winning coalitions of the bicameral standard procedure require the consent of the Commission and of the Council referring to unanimity, simple or qualified majority subgames of member states. The semi-tricameral cooperation procedure includes the EP in EU legislation in one out of two sets of feasible winning coalitions: the first set encompasses the Commission and all member states, the second set consists of coalitions comprising the Commission, more than 62 Council votes and at least half of the EP votes. The latter set of winning coalitions is also feasible under codecision procedure, but in this case the second set combines the unanimous member states with at least the absolute majority of EP votes. Since the Commission no longer has the right to withdraw its proposal when Council and Parliament conciliate their views in the second reading of the codecision procedure, the Commission can be excluded. Hence, under codecision procedure the EP holds the same position as the Commission under cooperation procedure. In this respect, both combinations of the two sets of winning coalitions install a semi-tricameral system: either the EP or the Commission can be excluded from EU legislation.

Legislative sets of winning coalitions represent the cornerstone of our analysis of EU legislative entities' decisiveness and inclusiveness. With regard to the fact that member states establish different legislative sets by introducing different procedures for EU policy areas, we investigate the reasons for member states making the choices for specific institutional settings that they do. We take into account the arguments of spatial analysts on the importance of actors' policy preferences by means of our inclusiveness index. In addition, we apply the relative decisiveness concept to the inter-institutional sets of winning coalitions in EU legislation. In the following section we argue that member states take into account the effects on both their decisiveness and their inclusiveness when they introduce or change the procedural settings for EU policy areas.

3 Decisiveness and Inclusiveness in European Legislation

The bicameral setting of the standard procedure between the Commission and the Council and the semi-tricameral participation of the EP under cooperation and of the Commission under codecision procedure suggest that member states try to reach different goals by Treaty reform, such as reducing EU transaction costs or decreasing the so-called democratic deficit (see Wessels 1991; Ludlow 1991). In the past, Treaty reforms have given the Commission functions of legislative agenda setting and safe-guarding, and the EP more rights in EU legislation. However, since the member states are the signatories of the EU constitution, we argue that their expected gains are the driving force behind the material integration of policy areas and the constitutional choice of different procedures. Thus, by focusing solely on the impact on qualified majority rule in the Council, many intergovernmental analyses are unable to provide insight on the reasons for institutional delegation.

This shortcoming is best illustrated by some of the partly striking, then again partly insufficient conclusions drawn from such voting power calculations. The most prominent result was the discovery of the "dummy player-position" of Luxembourg. According to relative voting power analysts, Luxembourg therefore did not realise the fact that it would have no relative power during the first EU Treaty era under qualified majority rule (Brams 1976). Second to this, Council power index analysis recently claimed to have "uncovered" the unfavourable British attitude towards the blocking minority rule, as the proposed increase from 23 to 26 minimum votes would reduce the British relative power share (Johnston 1995). Others argue that, due to the accession of new members, the relative decisiveness differences between unanimity and majority voting rules become less and less pronounced (Lane et al. 1995). Such striking conclusions prompt the question of whether member states actually misperceive the impact of Treaty reform or whether the study of relative decisiveness is an insufficient tool for explaining intergovernmental choice of EU voting rules (Garrett et al. 1995).

On closer inspection, indices on *relative decisiveness* are calculated using the concept of simple games with two properties: First, simple games only differentiate between winning and losing coalitions; and, second, they satisfy monotonicity assuming the continuance of a winning coalition in cases of additional members.[2] In the case of simple games, indices of relative decisiveness are single valued solution concepts on pivotal entities. Being pivotal can be interpreted as being a relative resource resulting from the entities' probability of realising their preference in the collective outcome. If member states have different voting weights under majority rule, these resources can be distributed asymmetrically. However, since member states make their constitutional decision on Treaty reform under unanimity, the question is why a member state should accept a higher voting weight of another member state providing the latter with more relative resources for future majority decision making.

In our view the constitutional choice of EU voting rules depends on the expected gains from potential legislation which are determined by both the number of feasible decisions and the distribution of their gains. When reforming the EU framework, member states' central motive is to improve their gains from future legislation based on their expected profits minus their expected costs of potential EU legislation. According to Buchanan and Tullock (1962, p. 70), signatories decrease the voting quota when all incumbents expect higher gains from future majority legislation. Accordingly, if member states expect to be affected similarly by future legislation, they establish the "One-Man-One-Vote" provision providing for a uniform distribution of expected gains. Consequently, different voting weights are established to obtain a balanced distribution of EU legislative gains if the status quo or the expected decision affect member states differently. For example, the unification of Germany had no effect on the distribution of member states' voting weights because the latter serve as a parameter for the distribution of expected gains rather than for the representative size of the member states' population.

Voting weights, minority blocking rules, veto player positions or multi-cameralism with different subgames are all methods of balancing the distribution of expected gains. Despite their procedural variety, all these methods may differentiate between the entities' relative ability of being decisive on any EU legislative proposal. Though relative power index analysis is widely used, its application on EU inter-institutional sets of winning coalitions imposes severe demands on the method of measurement. Compared to unicameral analysis, the normalisation over all entities must appropriately reflect the conditions for the different levels, the basic games, the subgames and the inter-institutional compound game (König and Bräuninger 1996, p. 338). Taking this into consideration, the most applicable concepts for the analysis of the relative decisiveness of entities in inter-institutional sets of winning coalitions are arguably the normalised Banzhaf and the Shapley-Shubik index (Nurmi 1987).

---

[2] For any coalition S of the actor set N, $v(S) = 1$ if S is winning, and $v(S) = 0$ if S is losing, where v represents the characteristic function; v is monotonic if $v(S) \geq v(T)$ for any $S \supseteq T$.

Although both indices have certain theoretical parallels, they differ with respect to their conceptions of critical defections. An entity's relative contribution to transforming winning into losing coalitions determines the relative Banzhaf power (Banzhaf 1965). In particular for inter-institutional sets of winning coalitions, the additivity of these critical positions must be called into question, since the Banzhaf index takes into account several critical positions in one single winning coalition (Dubey and Shapley 1979). This raises the question of how to interpret Banzhaf additive power, because highly vulnerable minimal winning coalitions become more important for the power calculation than those only made vulnerable by a few members (Shelley 1986). For this reason, the inter-institutional relationship of Banzhaf decisiveness is highly distorted by the different membership size of EU voting bodies. The Shapley-Shubik index refers to all possible voting sequences and checks how often each entity is able to transform a losing into a winning coalition (Shapley and Shubik 1954). An entity's decisiveness is defined as the probability of being pivotal, i.e. decisive in one of all equal probable voting sequences. Based on this probability concept, the individual Shapley-Shubik shares, ϕ_i, can be added up over any set of actors and be interpreted as an additive measure for relative coalitional power. We therefore apply the Shapley-Shubik index to measure individual decisiveness.

We regard relative decisiveness as being one major aspect of member states' constitutional choice. Following the same line of thought, we consider their choice of the strength of a voting rule to be the second major aspect of EU institutional integration because it influences the likely policy outcomes that will ensue. Weak voting rules, like simple majority, increase the number of feasible decisions by facilitating the possible exclusion of entities from the EU legislative set of winning coalitions, whereas unanimity guarantees high *inclusiveness* for all actors, resulting in a high status quo bias of single favourable winning coalitions. The rationale for member states' choice of unanimity rule might therefore be an expectation of low legislative gains either because of low profits or high costs. Accordingly, member states only expose themselves to the danger of exclusion if they expect higher profits from future EU legislation.

The strength of a voting rule refers to the entities' chances of being included in any potential decision. Since we assume Yes- and No-votes to have the same probability, all feasible coalitions are equiprobable. In simple games, the probability of an entity's inclusion varies between 0.5 and 1.0. Strong voting rules guarantee the inclusion of an entity's preferences in the collective decision, whereas the inclusion of an entity's preference is determined by luck if it can be excluded from any feasible winning coalition (Barry 1989, p. 287). Thus, the inclusiveness of a dummy player is still 0.5. Assuming v to be a simple game, where $v(S) = 1$ if S is winning, we define the inclusiveness index ω of actor i in the game v as

$$\omega_{\mathrm{i}}(v) = \frac{\sum\limits_{\mathrm{S} \subseteq \mathrm{N},\ \mathrm{i} \in \mathrm{S}} v(\mathrm{S})}{\sum\limits_{\mathrm{S} \subseteq \mathrm{N}} v(S)},$$

Fig. 2 Characterisation of voting procedures

		Decisiveness ϕ	
		Equal	*Unequal*
Inclusiveness ω *of an actor*	*high*	Unanimity (veto rights for all actors)	veto right for a single actor
	low	Unweighted majority voting	Weighted majority voting

i.e. i's number of participations in winning coalitions in relation to the number of all feasible winning coalitions (Bräuninger 1996).

Neither the relative nor the absolute aspect of voting power solely describe the choice of voting rules. In our view, voting rules are instruments that can be used to obtain a uniform distribution of legislative gains over all member states. Unanimity and majority voting rules steer the power to act by defining the number of feasible decisions, whereas various voting prerogatives, such as voting weights or single veto player positions, determine member states' chances to influence the outcome. Thus, only the combination of both aspects (in-)equality and strength, can offer a satisfactory account for the member states' choice of EU voting rules.

Figure 2 combines the instruments measured by relative decisiveness ϕ and absolute inclusiveness ω of member states. For the study of specific constitutional choices, we take into account both aspects of member states' expectations of potential EU legislation. Accordingly, the choice of the strength of voting rules depends on the expected gains determined by the number of feasible decisions, whilst the distribution is regulated by equal or unequal settings. The member states' expectation of a few decisions by uniform distribution of EU legislative gains favours the setting of unanimity, whereas a higher number by uniform distribution results in unweighted majority voting. Member states may also agree on single veto player positions when they expect a low number of decisions but an asymmetric distribution of EU legislative gains. Finally, weighted votes may be introduced in the case of a higher number of decisions by asymmetric distribution.

Although our scheme considers the different voting rules within the Council, the question of the participation of supranational entities still remains. Studying the inter-action between the Council, the Commission and the EP, recent applications of spatial models assumed extreme policy positions of supranational entities when determining the different procedural win sets (Steunenberg 1994; Tsebelis 1994; Schneider 1995). Under this assumption, the participation of the Commission and the EP may bias member states' legislative gains, prompting the question as to why some member states should accept the restriction of their own legislative profits. Leaving aside the assumption of extreme policy positions of supranational entities, we argue that the Commission and the EP are expected to increase the gains of the member states by promising to reduce both transaction costs and the charge democratic deficit. Since different procedures exist for EU legislation, the application of decisiveness and inclusiveness provides an insight into the member states' expectations of different policy areas.

4 Member States' Expectations of Policy Area Legislation

The consequences of different provisions for the Commission, the member states and the political groups in the EP are listed in Table 1. For the reasons discussed, we measure relative decisiveness by means of the Shapley-Shubik index ϕ and absolute inclusiveness by means of our index ω defined above. In the rows of Table 1 we list the entities grouped along EU chambers. The columns refer to our procedures and three different rules which may be applied to the standard procedure. Each of the six procedural settings has distinct effects on entities' decisiveness and inclusiveness.

Under standard procedure, we find equal and unequal settings with varying voting quotas. Although decisiveness ϕ does not differentiate between the member states either in the case of unanimity (0.0625) or simple majority (0.0333), their degrees of inclusiveness ω reveal the greatest difference. Unanimity guarantees the inclusion of all member states' policy preferences indicated by their maximal inclusiveness of 1.0. In the case of simple majority, by contrast, the danger of being excluded is very high (0.5500) approaching the dummy player's inclusion probability of 0.5. All member states, however, have the same absolute and relative power on different levels. Under qualified majority in standard, Article 148, 2b, cooperation and co decision procedure the inclusion probability ω of the four large member states is 86 and 85 %, respectively, while Luxembourg's inclusiveness ω varies between 57 and 61 %. The relative decisiveness of large member states is also higher here than in cases of equal settings. Qualified majority thus stresses the differences between the member states with regard to relative decisiveness and absolute inclusiveness.

Concerning the inter-institutional interaction in the standard procedure, the EP is a dummy player and can be excluded from building any feasible winning coalition. Hence, the EP cannot influence the outcome and its policy preference is included only by luck. The feature of the bicameral setting is illustrated by the Commission's inclusiveness ω and decisiveness ϕ. Under standard procedure the Commission's policy preference must be included in any legislative proposal, but its ability of being decisive varies widely. The Commission is an equal counterpart to all member states in cases of simple majority voting, but its decisiveness ϕ decreases from majority voting to unanimity. Hence, if the member states take a unanimous decision, the Commission has the lowest share of relative power. Except for the unicameral procedure of Article 148, 2b, the Commission's policy preferences are included in all EU legislation. Qualified majority discriminates between the member states, and the additional provision for a minority rule (Article 148, 2b)—the only unicameral procedure—not only favours the smaller member states' relative decisiveness ϕ but also increases their absolute inclusiveness ω.

Compared to qualified majority under standard procedure, the cooperation and co decision procedures have little effect on member states' inclusiveness ω. Only their decisiveness ϕ is modified as a result of the participation of the EP. However, the parliamentary entities' probability of being included in potential EU legislation

Table 1 Relative decisiveness (Shapley-Shubik ϕ_i) and absolute inclusiveness (ω_i) of EU legislative actors (August 1995)

	Votes	*Standard procedure*						*Article 148,2b*		*Cooperation*		*Codecision*	
		Unanimity		*Simple Majority*		*Qualified Majority*							
		ϕ_i	ω_i	ϕ_i	ω_i	ϕ_i	ω_i	ϕ_i	ω_i	ϕ_i	ω_i	ϕ_i	ω_i
Commission	1	0.0625	1.0000	0.5000	1.0000	0.3103	1.0000	0	0.5000	0.2841	1.0000	0.2216	0.9996
Council													
• France	10	0.0625	1.0000	0.0333	0.5500	0.0810	0.8627	0.1114	0.8490	0.0674	0.8627	0.0709	0.8627
• Germany	10	0.0625	1.0000	0.0333	0.5500	0.0810	0.8627	0.1114	0.8490	0.0674	0.8627	0.0709	0.8627
• Italy	10	0.0625	1.0000	0.0333	0.5500	0.0810	0.8627	0.1114	0.8490	0.0674	0.8627	0.0709	0.8627
• United Kingdom	10	0.0625	1.0000	0.0333	0.5500	0.0810	0.8627	0.1114	0.8490	0.0674	0.8627	0.0709	0.8627
• Spain	8	0.0625	1.0000	0.0333	0.5500	0.0662	0.8003	0.0920	0.7939	0.0552	0.8004	0.0587	0.8004
• Belgium	5	0.0625	1.0000	0.0333	0.5500	0.0377	0.6909	0.0563	0.7098	0.0312	0.6910	0.0347	0.6910
• Greece	5	0.0625	1.0000	0.0333	0.5500	0.0377	0.6909	0.0563	0.7098	0.0312	0.6910	0.0347	0.6910
• Netherlands	5	0.0625	1.0000	0.0333	0.5500	0.0377	0.6909	0.0563	0.7098	0.0312	0.6910	0.0347	0.6910
• Portugal	5	0.0625	1.0000	0.0333	0.5500	0.0377	0.6909	0.0563	0.7098	0.0312	0.6910	0.0347	0.6910
• Austria	4	0.0625	1.0000	0.0333	0.5500	0.0310	0.6556	0.0476	0.6798	0.0258	0.6557	0.0293	0.6557
• Sweden	4	0.0625	1.0000	0.0333	0.5500	0.0310	0.6556	0.0476	0.6798	0.0258	0.6557	0.0293	0.6557
• Denmark	3	0.0625	1.0000	0.0333	0.5500	0.0242	0.6167	0.0389	0.6486	0.0204	0.6169	0.0239	0.6169
• Ireland	3	0.0625	1.0000	0.0333	0.5500	0.0242	0.6167	0.0389	0.6486	0.0204	0.6169	0.0239	0.6169
• Finland	3	0.0625	1.0000	0.0333	0.5500	0.0242	0.6167	0.0389	0.6486	0.0204	0.6169	0.0239	0.6169
• Luxembourg	2	0.0625	1.0000	0.0333	0.5500	0.0141	0.5736	0.0251	0.6069	0.0120	0.5737	0.0155	0.5737
Sum of council	87	0.9375	–	0.5000	–	0.6897	–	1.0000	–	0.5744	–	0.6269	–
EP													
• Socialists	221	0	0.5000	0	0.5000	0	0.5000	0	0.5000	0.0551	0.8450	0.0596	0.8453
• European People's Party	172	0	0.5000	0	0.5000	0	0.5000	0	0.5000	0.0327	0.6529	0.0362	0.6530
• United Left	31	0	0.5000	0	0.5000	0	0.5000	0	0.5000	.0070	0.5552	0.0073	0.5552
• Liberal Dem. and Reformists	52	0	0.5000	0	0.5000	0	0.5000	0	0.5000	0.0126	0.6017	0.0132	0.6018
• Democratic Alliance	56	0	0.5000	0	0.5000	0	0.5000	0	0.5000	0.0140	0.6137	0.0146	0.6137

(continued)

Table 1 (continued)

	Votes	Standard procedure						Article 148,2b		Cooperation		Codecision	
		Unanimity		Simple Majority		Qualified Majority							
		ϕ_i	ω_i	ϕ_i	ω_i	ϕ_i	ω_i	ϕ_i	ω_i	ϕ_i	ω_i	ϕ_i	ω_i
• Radical Alliance	19	0	0.5000	0	0.5000	0	0.5000	0	0.5000	0.0040	0.5338	0.0042	0.5338
• Greens	25	0	0.5000	0	0.5000	0	0.5000	0	0.5000	0.0056	0.5469	0.0058	0.5470
• Europe of Nations	19	0	0.5000	0	0.5000	0	0.5000	0	0.5000	0.0040	0.5338	0.0042	0.5338
• FPÖ[a]	5	0	0.5000	0	0.5000	0	0.5000	0	0.5000	0.0010	0.5100	0.0010	0.5100
• Vlaams Blok[a]	2	0	0.5000	0	0.5000	0	0.5000	0	0.5000	0.0003	0.5033	0.0003	0.5033
• Front National Belgium[a]	1	0	0.5000	0	0.5000	0	0.5000	0	0.5000	0.0002	0.5017	0.0002	0.5017
• Front National France[a]	11	0	0.5000	0	0.5000	0	0.5000	0	0.5000	0.0024	0.5212	0.0025	0.5212
• Democratic Unionist[a]	1	0	0.5000	0	0.5000	0	0.5000	0	0.5000	0.0002	0.5017	0.0002	0.5017
• Alleanza Nazionale[a]	11	0	0.5000	0	0.5000	0	0.5000	0	0.5000	0.0024	0.5212	0.0025	0.5212
Sum of EP	626	0	–	0	–	0	–	0	–	.1415	–	0.1515	–

[a] Non-attached members

increases substantially. Introducing the EP as a third collective actor is thus an instrument geared towards including another dimension into EU legislation without increasing the member states' probability of having their preferences disregarded. Comparing the cooperation and the co decision procedure, the latter strengthens the decisive role of the Council in particular.

Finally, the combination of relative decisiveness ϕ and absolute inclusiveness ω gives a satisfactory account for the member states' choice of institutional settings when they expect legislative gains from potential EU legislation in specific policy areas. Although the participation of supranational entities, such as the Commission or the EP, may promise higher gains, the member states' expectation of potential EU legislative costs prohibits the material integration of further policy areas. Material integration is thus a function of the expected effects of institutional settings.

The selective application of procedural settings to EU policy areas may serve as an indicator for the specific gains member states expect from EU legislation. Not only do EU voting rules vary in the degree of inclusiveness and decisiveness, but even more to the point the provisions for EU legislation have been changed quite differently and discriminate even within policy areas. Table 2 lists the proportion of procedural settings for all EU policy areas that came into operation with the Treaty of Rome in 1958, the Single European Act in 1987 and the Maastricht Treaty in 1993. As the table indicates, the policy areas of agriculture, trade, association, institutional and final provisions have been excepted from constitutional modifications. Changes of the status quo in the areas of association, institutional and final provisions concern the core of the EU framework. When negotiating on the Rome Treaty, member states being in fear of many (unfavourable) decisions therefore preferred unanimity as the principle voting rule in these fields. By contrast, the policy areas of agriculture and trade are dominated by the provision of qualified majority rules with voting weights under standard procedure. According to our two aspects of constitutional choice, member states originally expected an asymmetric distribution of a higher number of EU decisions with additional gains by the Commission's role in reducing transaction costs. For both policy areas, characterised by the highest numbers of proposals and adoptions (König 1997, p. 86), member states have abstained from reducing the democratic deficit by excluding the EP.

In comparison, numerous modifications have been made in the areas of free movement, traffic, common rules and social policy which encompass the participation of the EP. The introduction of the cooperation procedure has also contributed to the reduction of the proportion of qualified and unanimous provisions. We observe a similar pattern for the introduction of the codecision procedure. Except for environmental policies, the co decision procedure has replaced the former provision for the cooperation procedure. Again, the recent introduction of industry policy does not promise EU legislative gains by a high number of decisions which would pave the way for weaker voting rules, whereas other areas introduced by the Maastricht Treaty provide for qualified majorities. In sum, different procedures and different voting rules regulate most EU policy areas. Our findings show a tendency towards weighted qualified majority voting in the Council either by modifications to the

Table 2 Proportion of procedural settings by treaty eras (%)

EU policy area	*Treaty of Rome (1958–1987)*			*Single European act (1987–1993)*				*Maastricht treaty (1993–)*				
	U	*Q*	*S*	*U*	*Q*	*S*	*CO*	*U*	*Q*	*S*	*CO*	*CD*
Principles/citizenship	60	40		43	43		14	67	22			11
Free movement of goods	22	78		12	88			12	88			
Agriculture	17	83		17	83			17	83			
Free movement of pers., serv., Cap	47	53		33	39		28	33	43			24
Transport	67	33		50	50			50	17		33	
Common rules	43	57		30	50		20	31	54			15
Economic policy	20	80		20	80			17	58		25	
Trade		100			100				100			
Social policy	40	40	20	33	33	17	17	25			50	25
Culture												100
Public health												100
Consumer protection												100
Transeuropean networks											50	50
Industry policy								100				
Economic and social cohesion				25	50		25	40	20		40	
Research and technical development				50			50	25	25		25	25
Environment				100				33			33	33
Development											100	
Association	100			100				100				
Institutional provisions	100			100				100				
Financial provisions	50	50		45	55			33	67			
Final provisions	80	20		80	20			80	20			
Sum per era (100 %)	51	48	1	45	45	1	9	38	38		12	11

U—Unanimity, Q—Qualified Majority, S—Simple Majority under Standard Procedure; CO—Cooperation Procedure; CD—Codecision Procedure
Source Compilation of own data, see König (1997)

standard procedure or by the introduction of the cooperation and codecision procedures. Despite this overall tendency, the member states have increased the proportion of unanimous voting rules in some policy areas, namely in the areas of the common rules and economic and social cohesion.

5 Conclusion

Looking beyond the scope of the analysis here, the Maastricht Treaty has brought about a new pattern of EU institutional integration. This new form of integration describes the move to selective expectations of potential legislative costs. It can be observed in the recent trend of including provisions for "opt-out" clauses as often favoured by either the United Kingdom or Denmark. The tendency towards this new pattern of selective EU integration has been reinforced in the provisions laid down for Monetary Union, as is fittingly illustrated by the current debate on the economic criteria for membership. Although enlargement by Eastern and Southern countries has rekindled the debate on core-membership, recent constitutional development has been characterised by the constitutional choice of voting rules applicable to all member states.

For the analysis of recent EU constitutional development we presented our approach on constitutional actors' expectation of legislative gains that could be obtained by the introduction or change of voting rules. In our view, the impact of voting rules on future decision making can be expressed by two aspects, the strength and the (in-)equality of their settings. Due to the fact that relative voting power analyses cannot consider the strength of voting rules, we introduced our concept of inclusive-ness measuring the frequency with which an actor will participate in winning coalitions in relation to the number of all winning coalitions. Since high inclusiveness of all actors results in high status quo probability, inclusiveness directly addresses one aspect of legislative gains, namely the number of feasible coalitions in which an actor can realise his preference.

However, the second aspect of legislative gains concerns their distribution among actors. In order to steer the distribution of expected legislative gains, constitutional actors may establish either equal or unequal settings, the latter privileging some actors by providing different voting weights or actor-specific veto rights. Their effects can be measured by means of the Shapley-Shubik index which calculates actors' relative abilities of being decisive in forming winning coalitions. In the case of the equal "One-Man-One-Vote" provision, actors are provided with the same relative ability to influence the distribution of expected legislative gains, while unequal provisions introduce actor-specific prerogatives.

The reason for the establishment of unequal settings might be that all constitutional actors agree to balance the distribution of gains when certain actors are considered to have a higher status quo bias. In the past, constitutional actors favoured a common solution rather than allowing for core-membership. Core-membership, however, has already been applied to EU social politics and Monetary Union. According to our approach, there might be two reasons for core-membership: either EU core-member-ship provides for even higher expected gains or constitutional actors could not agree on a formula for balancing the distribution of gains.

References

Achen, C.H. (1995). *How Can We Tell a Unitary Rational Actor when We See One?* Paper prepared for presentation at the annual meeting of the Midwest Political Science Association, Chicago.

Attina, F. (1990). The voting behavior of the European Parliament Members and the problem of Europarties. *European Journal of Political Research, 5*, 557–579.

Barry, B. (1989). *Theories of justice: a treatise on social justice*. London: Harvester-Wheatsheaf.

Banzhaf, J. F. (1965). Weighted voting doesn't work: a mathematical analysis. *Rutgers Law Review, 19*, 317–343.

Brams, S. J. (1975). *Game theory and politics*. New York: Free Press.

Brams, S. J. (1976). *Paradoxes in Politics*. New York: Free Press.

Brams, S. J., & Affuso, P. (1985). New paradoxes of voting power on the EC council of ministers. *Electoral Studies, 4*, 135–139.

Bräuninger, T. (1996), *Die Modellierung von Entscheidungsverfahren internationaler Organisationen am Beispiel der Meeresbodenbehörde. Teilnahme-, Mitwirkungs- und Durchsetzungschancen in einem institutionalisierten Regime (unpublished thesis)*, Mannheim: University of Mannheim.

Buchanan, J. M., & Tullock, G. (1962). *The calculus of consent: logical foundations of constitutional democracy*. Ann Arbor: University of Michigan Press.

Coleman, J. S. (1971). Control of collectivities and the power of a collectivity to act. In B. Lieberman (Ed.), *Social choice* (pp. 269–299). New York: Gordon and Breach.

Dubey, P., & Shapley, L. S. (1979). Mathematical properties of the Banzhaf power index. *Mathematics of Operations Research, 4*, 99–131.

Fischer, D., & Schotter, A. (1978). The inevitability of the paradox of redistribution in the allocation of voting weights. *Public Choice, 33*, 49–67.

Garrett, G., & Tsebelis, G. (1996). an institutional critique of intergovernmentalism. *International Organization, 50*, 269–299.

Garrett, G., McLean, I., & Machover, M. (1995). Power, power indices and blocking power: a comment on Johnston. *British Journal of Political Science, 25*, 563–568.

Hosli, M. O. (1993). Admission of European free trade association states to the European community: effects on voting power in the European community council of ministers. *International Organization, 47*, 629–643.

Jacobs, F., Corbett, R., & Shackleton, M. (1992). *The European Parliament*. London: Longman.

Johnston, M. (1994). *The European council*. Oxford: Westview.

Johnston, R. J. (1995). The conflict over qualified majority voting in the European Union Council of ministers: an analysis of the uk negotiating stance using power indices. *British Journal of Political Science, 25*, 245–288.

Koehler, D. H. (1990). The size of the Yolk: computations of odd and even-numbered committees. *Social Choice and Welfare, 7*, 231–245.

König, T. (1997). *Europa auf dem Weg zum Mehrheitssystem? Gründe und Konsequenzen nationaler und parlamentarischer Integration*. Opladen: Westdeutscher Verlag.

König, T., & Bräuninger, T. (1996). Power and political coordination in america and german multi-chamber legislation. *Journal of Theoretical Politics, 8*, 331–360.

König, T., & Bräuninger, T. (1998). The inclusiveness of European decision rules. *Journal of Theoretical Politics, 10*, 125–141.

Lane, J.-E., Mæland, R., & Berg, S. (1995). The EU parliament: seats, states and political parties. *Journal of Theoretical Politics, 7*, 395–400.

Ludlow, P. (1991). The European Commission. In R. O. Keohane & S. Hoffman (Eds.), *The new European community: decision making and institutional change* (pp. 85–132). Boulder, Col.: Westview.

McKelvey, R. (1979). General conditions for global intransitivities in formal voting models. *Econometrica, 47*, 1085–1112.

Nurmi, H. (1987). *Comparing voting systems*. Dordrecht: Reidel Publishing.

Schneider, G. (1995). The limits of self-reform: institution-building in the European Union. *European Journal of International Relations, 1*, 59–86.

Shapley, L. S., & Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review, 48*, 787–792.

Shelley, F. M. (1986). Voting power in municipal annexation elections. *Quality and Quantity, 120*, 257–260.

Spence, D. (1994). Structure, functions and procedures in the commission. In G. Edwards & D. Spence (Eds.), *The European Commission* (pp. 92–116). Essex: Longman.

Steunenberg, B. (1994). Decision making under different institutional arrangements: legislation by the European community. *Journal of Institutional and Theoretical Economics, 150*, 642–669.

Tsebelis, G. (1994). The power of the European Parliament as a conditional agenda setter. *American Political Science Review, 88*, 128–142.

Usher, J. (1994). The commission and the law. In G. Edwards & D. Spence (Eds.), *The European ommission* (pp. 146–168). Essex: Longman.

Wessels, W. (1991). The EC council: the community's decision-making center. In R. O. Keohane & S. Hoffman (Eds.), *The new European Community: Decision-making and Institutional Change* (pp. 133–154). Boulder, Col.: Westview.

Westlake, M. (1994). *The commission and the parliament*. London: Butterworths.

Widgrén, M. (1994). Voting power in the EC decision-making and the consequences of two different enlargements. *European Economic Review, 38*, 1153–1170.

Minimax Multi-District Apportionments

Gianfranco Gambarelli and Arsen Palestini

1 Introduction

Apportionments are a typical problem of the world of politics, as there is a need to assign seats to parties in proportion to the number of votes, or constituencies in proportion to the population. The problem consists in transforming an ordered set of nonnegative integers, the “votes”, into a set of integers, the “seats”, respecting some specific fairness conditions. Several methods have been constructed, but paradoxes and contradictions are likely to occur in many cases [e.g., Brams (1976)]. Starting from some results by Balinski and Demange (1989) and Balinski and Young (1982), Gambarelli (1999) proposed an apportionment technique respecting the principal criteria of electoral systems. The approach was related to one-district elections and involves the determination of an order of preference for the satisfaction of criteria and introduces the concept of “minimax solution”.

In this chapter we propose a generalization of the above method to multi-district systems, where further problems arise. In such situations, the total apportionment depends not only on the global number of votes, but also on the votes obtained by parties in every district.

The basic apportionment criteria are presented in the following Section. Section 3 synthetically features the most known apportionment techniques. In Sect. 4 we recall the minimax method for one-district apportionments. The generalization of this method to the multi-district case is presented in Sect. 5. The ordering of the new criteria is discussed in Sect. 6. A theorem on the existence of the solutions is presented in Sect. 7. Section 8 shows further criteria to refine the

G. Gambarelli (✉)
Department of Management, Economics and Quantitative Methods,
University of Bergamo, Bergamo, Italy
e-mail: gambarex@unibg.it

A. Palestini
MEMOTEF, Sapienza University of Rome, Rome, Italy
e-mail: arsen.palestini@uniroma1.it

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_10,

solution. Section 9 supplies a comparison with the results of classical methods. An overview on the Banzhaf index and an algorithm generating solutions are reported in the Appendices.

2 Criteria

Apportionment can be defined as the process of allotting indivisible objects (seats) amongst players (parties) entitled to various shares. The related literature is quite vast: see for instance, Hodge and Klima (2005) for an overview.

Let $v = (v_1, \ldots, v_n)$ be the vector of valid votes won by the n parties ($n \geq 2$) of an electoral system, where s^T is the total number of seats to be assigned. We call:

$v^T = \sum_{i=1}^{n} v_i$ the total number of votes,

$h_i = v_i \cdot s^T / v^T$ the **Hare quota** of the i-th party,

S^O the set of n-dimensional integer allotments $s = (s_1, \ldots, s_n)$ such that

$$\sum_{i=1}^{n} s_i = s^T.$$

The problem consists in detecting, amongst all possible seat allotments, the aptest one to represent the Hare quota vector $\tilde{h} = (h_1, \ldots, h_n)$. Obviously, if $\tilde{h}$ is an integer vector, it turns out to be the best solution. Otherwise, a rounding allotment procedure is necessary, which should satisfy some fairness criteria. The most well known of these criteria are

a) **Hare maximum:** No party can obtain more seats than the ones it wins by rounding up its Hare quota.
b) **Hare minimum:** No party can obtain fewer seats than the ones it wins by rounding down its Hare quota.
c) **Monotonicity:** For any pair of parties, the one entitled to fewer votes cannot win more seats than the other one.
d) **Superadditivity:** A party formed by the union of two parties must at least gain a number of seats equal to the sum of the seats won by the single parties.
e) **Symmetry:** The apportionment must not depend on the order in which parties are considered. In particular, two parties having the same amount of votes must achieve the same number of seats.

It is well-known that no apportionment method exists which conjointly verifies all the above conditions, for all possible vectors of votes. For instance, consider a system in which there are only two parties gaining exactly the same amount of votes, and where an odd number of seats must be allotted. In such a case, criterion

e) cannot be fulfilled, then an exogenous criterion must be applied. Analogously, it can be proved that c) and e) cannot be respected in general, if a) and d) hold and vice versa. Moreover, some paradoxes may occur: “Alabama”, “Population”, “New States” and so on [e.g. Brams (1976)].

Hence, the problem we are going to face is the search for a suitable compromise solution.

3 Classical Apportionment Methods

Hereafter some traditional apportionment techniques will be recalled and applied to a simple apportionment problem.

Example 1 Let 19, 15, 6 be the valid votes obtained by the parties A, B, C, and 4 seats be shared among these parties.

The **Method of Largest Remainders** (also known as **Hamilton’s Method**) assigns the initial seats according to the Hare minimum quotas and the remaining ones to the parties having the largest fractional parts of their quotients among the remainders.

In this case, party A and party B initially win one seat each. Subsequently, party A with 0.9 and party C with 0.6 obtain the last two seats. Hence Hamilton’s Method provides the seat allotment (2, 1, 1).

The **Method of the Greatest Divisors** (also known as **Method of d’Hondt** or **Jefferson’s Method**) allots seats to the parties having the highest quotients after dividing their respective shares by 1, then by 2, then by 3, and so on. In our case, only the division by 2 is needed, because the quotients it generates are 9.5 for A, 7.5 for B, 3 for C. Consequently, the highest among all quotients are 19, 15, 9.5 and 7.5, so A and B gain two seats each, and no seat is assigned to C.

The **Method of the Greatest Divisors with quota** (also known as **Balinski-Young Method**) is an apportionment technique similar to that of d’Hondt, except for the impossibility for each party to exceed its Hare maximum quota: when a party reaches its Hare maximum quota, it does not participate in the seat allotment any longer. In this example, no party can exceed that quota, so the apportionments generated by the Method of d’Hondt and by Balinski-Young Method coincide.

For further apportionment techniques see for instance Nurmi (1982), Holubiec and Mercik (1994) and Hodge and Klima (2005).

4 The Method of Minimax

The minimax method is inspired by the nucleolus (see Schmeidler 1969).

4.1 Preliminary Definitions

Let s be a seat vector of $S^{\,0}$ and v be a vote vector in R^n_+. Consider the simplex $\bar{X} = \left\{(x_1, \ldots, x_n) \in R^n_+ : \sum_{k=1}^{n} x_k = 1\right\}$.

Given a transform $t : R^n_+ \to \bar{X}$ we call $t(s) = \bar{s} = (\bar{s}_1, \ldots, \bar{s}_n)$, $t(v) = \bar{v} = (\bar{v}_1, \ldots, \bar{v}_n)$.

For all $s \in S^{\,0}$, $v \in R^n_+$, $i, j = 1, \ldots, n$, $i \neq j$, we call

$e_j(v, s) = \bar{s}_j - \bar{v}_j$ the **bonus** of the j-th component;

$c_{ij}(v, s) = e_i(v, s) - e_j(v, s)$ the **complaint** of the j-th party against the i-th party;

$c(v, s)$ the **complaint vector**, i.e. the vector whose components are the non-negative complaints listed in non-increasing order.

The previous definitions allow us to establish a relation on S.

For all s', $s'' \in S^{\,0}$ we say that:

- s' is **indifferent** to s'' with respect to v ($s' \approx s''$) if and only if $c\,(v, s') = c\,(v, s'')$.
- s' is **preferable** to s'' with respect to v ($s' \succ s''$) if and only if an integer k exists such that:

1. $c_k(v, s') < c_k(v, s'')$;
2. $c_h(v, s') = c_h(v, s'')$ for all $h < k$.

It is easy to prove that $\approx$ is an equivalence relation and that $\succ$ is a total order in the set $S^{\,0}$. Consequently, this relation determines a preference for the apportionment vectors of $S^{\,0}$.

Observe that, if a transform $t^* : R^n_+ \to \bar{X}$ exists such that $t^*(s) = t^*(v)$, then all bonuses and consequently all parties' complaints vanish.

We call ***t*-minimax criterion** (or ***t*-criterion**) the criterion which consists in keeping only the seat allotments not preferred, with respect to the distribution of votes, by other apportionments, and discarding all the others.

Gambarelli's method (1999) consists in the following procedure.

An order of importance of criteria to be applied, is preliminarily fixed: C_1, C_2, …, C_k

Then we call:

$S^{\,1}$ the subset of $S^{\,0}$ obtained after applying criterion C_1;

$S^{\,2}$ the subset of $S^{\,1}$ obtained after applying criterion C_2; and so on until $S^{\,k}$.

We call C_1 C_2 … C_k**-solution** the set $S^{\,k}$ of allotments which respect the criteria C_1, C_2, …, C_k, applied in sequence.

The first criterion to be applied in this method is called the ***F*-criterion**, and consists in discarding all seat apportionments violating at least one of the basic criteria: Hare maximum, Hare minimum and Monotonicity. Gambarelli (1999, p. 446) proved that the F-criterion applied as the first criterion C_1 in the sequence of criteria determining the solution, generates a non-empty set of seat allotments.

4.2 The N-criterion

Let $(x_1, \ldots, x_n)$ be a nonnegative integer vector such that $\sum_{j=1}^{n} x_j = x^T$.

Consider the normalization map $N : R_+^n \to \bar{X}$ such that $N(x_1, \ldots, x_n) = \left(\frac{x_1}{x^T}, \ldots, \frac{x_n}{x^T}\right) = (\bar{x}_1, \ldots, \bar{x}_n)$.

The **Normalization criterion** (or the ***N*-criterion**) is the t-minimax criterion associated to the transform $t = N$.

Notice that the apportionments generated by the application of the N-criterion coincide with those provided by Hamilton's Method. Anyway, the next criterion to be applied will furthermore restrict the solution set achieved as yet.

4.3 The β-criterion

This criterion is based on the Banzhaf normalized power index (1965). Some notes on this index (here simply called "***β*-power index**") are supplied in Appendix 1. We consider the β-power index particularly suitable for electoral systems, because of its proportionality properties in the allotment of seats.

In order to enunciate the second minimax criterion, we will consider the transform $\bar{\beta} : R_+^n \to \bar{X}$, associating to every vote distribution $v = (v_1, \ldots, v_n)$ and to every seat allotment $s = (s_1, \ldots, s_n)$ the Banzhaf normalized power indices $\bar{\beta}(v)$ and $\bar{\beta}(s)$ of the related voting games with simple majority quota.

The ***β*-criterion** is the t-minimax criterion associated to the transform $t = \bar{\beta}$.

4.4 Minimax Solutions

The two previous minimax criteria allow us to make use of two different sequences of criteria: by choosing $C_1 = F$, $C_2 = N$, $C_3 = \beta$, we will have the $FN\beta$-solution; by choosing $C_1 = F$, $C_2 = \beta$, $C_3 = N$, the method will generate $F\beta N$-solution. Gambarelli (1999, p. 455) proved that each $FN\beta$-solution and each $F\beta N$-solution consist of a non-empty sets of seat allotments. The next example shows an application of the minimax method.

Example 2 Let 8 seats be assigned to the 4 parties A, B, C, D entitled to the votes (50, 30, 15, 5).

If we apply the F-criterion, then all seat apportionments are discarded except s_1 = (4,2,1,1), s_2 = (4,3,1,0) and s_3 = (4,2,2,0). In fact, Hare quotas are: 4 for party A, 2.4 for party B, 1.2 for party C, and 0.4 for party D, so all the remaining seat distributions would violate either Hare maximum or Hare minimum.

Subsequently, N-criterion is applied to the apportionment set: $S^1 = \{(4,2,1,1),(4,3,1,0),(4,2,2,0)\}$.

The normalized vector of votes is $\bar{v}$ = (0.5, 0.3, 0.15, 0.05)

The normalized vectors of the seat distributions are respectively:

$\bar{s}_1$ = (0.5, 0.25, 0.125, 0.125), $\bar{s}_2$ = (0.5, 0.375, 0.125, 0), $\bar{s}_3$ = (0.5, 0.25, 0.25, 0),

The bonuses are:

$e(v, s_1)$ = (0, −0.05, −0.025, 0.075),

$e(v, s_2)$ = (0, 0.075, −0.025, −0.05),

$e(v, s_3)$ = (0, −0.05, 0.1, −0.05).

Consequently, the three complaint vectors are:

$c(v, s_1)$ = (0.125, 0.100, 0.075, 0.050, 0.025, 0.025),

$c(v, s_2)$ = (0.125, 0.100, 0.075, 0.050, 0.025, 0.025),

$c(v, s_3)$ = (0.150, 0.150, 0.100, 0.050, 0.050, 0.000).

According to the previously defined relation, it is $s_1 \approx s_2 \succ s_3$.

The application of the N-criterion causes the elimination of s_3. Then the set of "surviving" allotments is $S^2 = \{(4,2,1,1),(4,3,1,0)\}$.

The next step is the application of the β-criterion. The Banzhaf normalized power index of votes for simple majority can be obtained after some computations: $\bar{\beta}(v)$ = (0.7, 0.1, 0.1, 0.1).

The β-index of seats of each apportionment, for simple majority, has to be computed for s_1 = (4, 2, 1, 1) and s_2 = (4, 3, 1, 0):

$\bar{\beta}(s_1)$ = (0.7, 0.1, 0.1, 0.1), $\bar{\beta}(s_2)$ = (0.6, 0.2, 0.2, 0).

The bonuses are respectively:

$\bar{\beta}(s_1) - \bar{\beta}(v) = (0, 0, 0, 0)$, $\bar{\beta}(s_2) - \bar{\beta}(v) = (-0.1, 0.1, 0.1, -0.1)$.

So this last criterion yields the unique $FN\beta$-solution $S^3 = \{(4,2,1,1)\}$.

In general, S^3 may be composed by more than one seat allotment.

5 Multi-District Apportionments

Our aim is to extend the minimax method to the multi-district case.

5.1 A Leading Example

We will show our model using the following

Example 3 An electoral system is composed of two districts (to which 6 and 5 seats must be assigned) and three parties A, B, C. The valid votes obtained are shown in Table 1.

The local Hare quotas are reported in Table 2. The last row of Table 2 shows the global Hare quotas, i.e. the Hare quotas of the totals of Table 1. Notice that the

Table 1 The votes of Example 3

Votes	Party A	Party B	Party C	Totals
District I	50	60	10	120
District II	10	10	60	80
Totals	60	70	70	200

Table 2 The local and global Hare quotas of Example 3

	A	B	C	Totals
District I	2.500	3.000	0.500	6
District II	0.625	0.625	3.750	5
Totals	3.125	3.625	4.250	11
Global Hare quotas	3.300	3.850	3.850	11

Table 3 The normalized votes and β-indices of votes of Example 3

	Local normalized votes				Local β-indices of votes			
	A	B	C	Totals	A	B	C	Totals
District I	$0.41\overline{6}$	0.50	$0.08\overline{3}$	1	1/5	3/5	1/5	1
District II	0.125	0.125	0.750	1	0	0	1	1
Global	0.30	0.35	0.35	1	1/3	1/3	1/3	1

sum of local Hare quotas differs from the global Hare quotas. Table 3 shows the related normalized votes and β-indices of votes. In this case the last row shows these data at the global level, too.

5.2 Data and Variables

We will utilize the following indices:

d to denote the districts $(d = 1, \ldots, n_d)$ and

p to denote the parties $(p = 1, \ldots, n_p)$.

A multi-district apportionment problem is based on the following data:

$V = [v_{dp}]$ the matrix of valid votes obtained by the p-th party in the d-th district
$a = [a_d]$ the vector of the seats to be assigned to the d-th district
[in our example (6, 5)].

Other computation parameters are:

$S = [s_{dp}]$ any integer matrix whose elements are seat distributions which respect the total seats to be allotted in the districts, i.e. $\sum_{p=1}^{n_p} s_{dp} = a_d$ $(d = 1, \ldots, n_d)$

$b = [b_p]$.the total seats assigned to the p-th party in matrix S, i.e. $\sum_{d=1}^{n_d} s_{dp} = b_p$ $(p = 1, \ldots, n_p)$.

Let S^{0} be the set of matrices S respecting the above conditions.

We will generalize the definitions of Sect. 4.1 as follows.

For all $v \in V$, $S \in S^0$, p, $q = 1, \ldots, n_p$, $p \neq q$, we call

$e_{dp}(V, S) = \bar{s}_{dp} - \bar{v}_{dp}$ the **bonus** of the p-th party in the d-th district;

$c_{dpq}(V, S) = e_{dp}(V, S) - e_{dq}(V, S)$ the **complaint** of the p-th party against the q-th partyin the d-th district;

$c(V, S) = [c_1(V, S), \ldots, c_k(V, S)]$ the ***S*-complaint vector**, i.e. the vector whose components are the non-negative complaints of the whole matrix S, listed in non-increasing order.

The above definitions allow us to establish on S^O the same preference relationship introduced in Sect. 4.1.

5.3 Solutions

We will use the same concept of solution introduced in Sect. 4.1, with the simple substitution of S^k with S^k for all involved k.

For an easier understanding of the criteria used, we will present them together with the construction of the solution to Example 3. Obviously, the order of criteria can be changed depending on the importance given to them. Here the following sequence is used:

F_G**-criterion** (F-criterion for the global apportionments);

N_G**-criterion** (N-criterion for the global apportionments);

β_G**-criterion** (β-criterion for the global apportionments);

N_L**-criterion** (N-criterion for the local apportionments).

β_L**-criterion** (β-criterion for the local apportionments);

The F_G-**criterion,** N_G**-criterion** and β_G**-criterion** are no other than the corresponding criteria presented in Sect. 4.1, applied to global Hare quotas of the votes. In our example the application of the F_G-criterion leads to the only matrices where total seats per party are: (3, 4, 4), inasmuch this distribution is the only one which respects monotonicity, the Hare minimum and Hare maximum at global level. As the F_G-**criterion** supplies only one allocation of total seats, the N_G-criterion and the β_G-criterion maintain the set of the above matrices unchanged.

The N_L**-criterion** consists in keeping only the matrices which minimize the S-complaint vectors, according to what is indicated in the t-minimax criterion presented in Sect. 4.1, using $t = N$. In our example, to help the search for such matrices, we can focus on the only ones that respect the Hare minimum and Hare maximum in all the districts, as they are preferable to all the others. These are shown in the upper part of Table 4. In the same table the rounded normalized seats $\bar{s}_{dp}$, the bonuses e_{dp}, the complaints c_{dpq} and the S-complaint vectors c are shown.

The maximum values of the S-complaint vectors of the four matrices are respectively 0.20, 0.25, 0.225, 0.425. The matrix which corresponds to the minimum of such values is the first. Then the $F_G N_G \beta_G N_L$-solution is unique and is the matrix shown in Table 5.

Table 4 The computations to obtain the $F_G N_G \beta_G N_L \beta_L$-solution of Example 3

S	3	3	0	2	4	0	2	3	1	1	4	1
	0	1	4	1	0	4	1	1	3	2	0	3
$\bar{s}_{1p}$	0.50	0.50	0.00	$0.\overline{3}$	$0.\overline{6}$	0.00	$0.\overline{3}$	0.50	$0.1\overline{6}$	$0.1\overline{6}$	$0.\overline{6}$	$0.1\overline{6}$
$\bar{s}_{2p}$	0.00	0.20	0.80	0.20	0.00	0.80	0.20	0.20	0.60	0.40	0.00	0.60
e_{1p}	$0.08\overline{3}$	0.000	$-0.08\overline{3}$	$-0.08\overline{3}$	$0.1\overline{6}$	$-0.08\overline{3}$	$-0.08\overline{3}$	0.000	$0.08\overline{3}$	−0.850	$0.1\overline{6}$	$0.08\overline{3}$
e_{2p}	−0.125	0.075	0.050	0.075	−0.125	0.050	0.075	0.075	−0.150	0.275	−0.125	−0.150
c_{1pq}	0.17	0.08	0.08	0.25	0.25	0.00	0.17	0.08	0.08	0.42	0.33	0.08
c_{2pq}	0.20	0.175	0.025	0.20	0.175	0.025	0.225	0.225	0.00	0.425	0.40	0.025
c	0.20	0.175	0.17	0.25	0.25	0.20	0.225	0.225	0.17	0.425	0.42	0.40
	0.08	0.08	0.025	0.175	0.225	0.00	0.08	0.08	0.00	0.33	0.08	0.025

Table 5 The $F_G N_G \beta_G N_L \beta_L$-solution of Example 3

3	3	0
0	1	4

Table 6 The $F_G N_G \beta_G \beta_L$-solution of Example 3

2	3	1	1	3	2
1	1	3	2	1	2

Table 7 The $F_G N_G \beta_G \beta_L N_L$-solution of Example 3

2	3	1
1	1	3

Table 8 The votes of Example 4

Votes	Party A	Party B	Totals
District I	320	80	400
District II	4,680	4,920	9,600
Totals	5,000	5,000	10,000

The β_L**-criterion** consists in keeping only those matrices which minimize the S-complaint vectors, according to what is indicated in the t-minimax criterion presented in Sect. 4.1, using $t = \beta$. In our example, due to uniqueness, the $F_G N_G \beta_G N_L \beta_L$-solution coincides with the $F_G N_G \beta_G N_L$-solution.

If we want to change the order of the two local criteria, we must apply the β_L-criterion to the matrices of the $F_G N_G \beta_G$-solution. Observe that, among such matrices, there are two, and only two, which perfectly respect the power indices of the votes, shown in Table 3. These matrices (shown in Table 6) lead to null S-complaint vectors and therefore are the $F_G N_G \beta_G \beta_L$-solution. It is now easy to verify that the $F_G N_G \beta_G \beta_L N_L$-solution is the one shown in Table 7.

Observe that all the above solutions remain the same if the order of N_G-criterion and β_G-criterion is exchanged, as mentioned at the beginning of the presentation of these criteria. However, the two solutions obtained by inverting the order of local criteria are different. An example is now given in which these solutions coincide.

Example 4 An electoral system is made up of two districts (to which 20 and 80 seats must be assigned) and two parties A, B. The valid votes obtained are shown in Table 8.

The local and global Hare quotas are reported in Table 9. Observe that all Hare quotas are integer numbers. Table 10 shows the related normalized votes and β-indices of votes.

It is easy to verify that the $F_G N_G$-solution is the set of matrices shown in Table 11, varying the integer k from 0 to 20. Similarly for the $F_G \beta_G$-solution.

Table 9 The local and global Hare quotas of Example 4

Districts	A	B	Totals
I	16	4	20
II	39	41	80
Totals	55	45	100
Global Hare quotas	50	50	100

Table 10 The normalized votes and β-indices of votes of Example 4

	Local normalized votes			Local β-indices of votes		
Districts	A	B	Totals	A	B	Totals
I	0.8000	0.2000	1	1	0	1
II	0.4875	0.5125	1	0	1	1
Global	0.5	0.5	1	1/2	1/2	1

Table 11 The solutions of Example 4

Districts	A	B	Totals
I	$20 - k$	k	20
II	$30 + k$	$50 - k$	80
Totals	50	50	100

Table 12 The computations to obtain the $F_G N_G \beta_G N_L \beta_L$-solution of Example 4

$\bar{s}_{1p}$	$(20 - k)/20$	$k/20$
$\bar{s}_{2p}$	$(30 + k)/80$	$(50 - k)/80$
e_{1p}	$(4 - k)/20$	$(k - 4)/20$
e_{2p}	$(9 - k)/80$	$(k - 9)/80$
c_{1pq}	$4 - k/10$	
c_{2pq}	$9 - k/40$	

Obviously, the $F_G N_G \beta_G$-solution and the $F_G \beta_G N_G$-solution coincide with the above solutions.

Now we will continue with the calculations of the $F_G N_G \beta_G N_L$-solution (see Table 12).

After some algebra we obtain that 5 is the value of k which minimizes max $(4 - k/10, 9 - k/40)$. Therefore the $F_G N_G \beta_G N_L$-solution is made up of the only matrix shown in Table 13 and coincides with the $F_G N_G \beta_G N_L \beta_L$-solution.

Going on to the calculation of the $F_G N_G \beta_G \beta_L N_L$-solution, it is easy to verify that such a solution is the set of matrices shown in Table 11 for which $20 - k > k$ and $30 + k < 50 - k$. These matrices correspond to the values of k between 0 and 9. $k = 5$ is included in these. Therefore the $F_G N_G \beta_G \beta_L N_L$-solution coincides with the $F_G N_G \beta_G N_L \beta_L$-solution.

Table 13 The $F_G N_G \beta_G N_L \beta_L$-solution and the $F_G N_G \beta_G \beta_L N_L$-solution of Example 4

15	5
35	45

6 On the Ordering of Criteria

In the examples in the last Section we gave greater importance to global level criteria than to those at a local level; however, there is no change in the technique if the order is permuted. However, it seems reasonable to apply the F_G-criterion first, as this guarantees respect to the will of the entire electorate. Complaints are often heard about the misrepresentations of parliamentary majorities, due to local roundings. Such dissatisfaction seems reasonable inasmuch as a Parliament represents the entire population. Subsequently, the choice of order of the criteria depends on the national situation which it is applied to. In particular, the choice of priority between adhering to normalized votes or to power indices in the first case gives preference to the proportional aspect; in the second case to the majority aspect, which is essential for democracy.

7 On the Existence of Solutions

Theorem *For every multi-district apportionment problem, all solutions having the F_G-criterion as the first criterion, are not empty.*

Proof In our hypotheses at least one distribution of total seats $b = (b_1, \ldots, b_{n_p})$ which are able to verify the F-criterion exists (see the end of Sect. 4.1). It is known that, given any two integer vectors $a = (a_1, \ldots, a_{n_d})$ and $b = (b_1, \ldots, b_{n_p})$ having equal sums of the components, at least one integer matrix $(n_d \times n_p)$ exists, which has such vectors as totals of row and column. Each one of the other criteria C_{k+1} generates a nonempty subset of the S^k-solution, inasmuch it chooses, from these matrices, only the optimal ones according to that criterion; however, in the case of equal optimality, it keeps them all. ■

An algorithm for the automatic computation of the solution is shown in Appendix 2.

8 Further Criteria

In cases of non-uniqueness, further criteria can be added and applied in order to restrict the solution set, using the same techniques. For instance, after the five criteria have been applied, it is again possible to choose whether to give preference to the β_L-criterion or the N_L-criterion. Taking into consideration the corresponding complaints, it is possible to keep only those matrices for which the maximum of such a

Table 14 The allocations assigned by various methods in the case of Example 3

Method	Hamilton	Hondt-Jefferson	Balinski-Young	$F_G N_G \beta_G N_L \beta_L$ and $F_G \beta_G N_G N_L \beta_L$	$F_G N_G \beta_G \beta_L N_L$ and $F_G \beta_G N_G \beta_L N_L$
District	A B C	A B C	A B C	A B C	A B C
I	3 3 0	3 3 0	3 3 0	3 3 0	2 3 1
II	1 0 4	0 0 5	0 1 4	0 1 4	1 1 3
Totals	4 3 4	3 3 5	3 4 4	3 4 4	3 4 4
Breaks					
Local					
Hare maximum		X			
Power index	X	X	X	X	
Global					
Symmetry	X				
Monotonicity	X				
Hare maximum		X			

Table 15 The allocations assigned by various methods in the case of Example 4

Method	Hamilton Hondt-J. Balinski-Y	$F_G N_G \beta_G N_L \beta_L$ $F_G \beta_G N_G N_L \beta_L$ $F_G N_G \beta_G \beta_L N_L$ $F_G \beta_G N_G \beta_L N_L$
District	A B	A B
I	16 4	15 5
II	39 41	35 45
Totals	55 45	50 50
Breaks		
Local		
Hare minimum		X
Hare maximum		X
Global		
Symmetry	X	
Hare minimum	X	
Hare maximum	X	
Power index	X	

vector corresponds to a minimum number of votes. Resorting to this method a further restriction of the solution set can be obtained. The uniqueness of the final matrix, however, cannot be guaranteed; e.g. in the case where the global Hare quotas are of the type shown in the example in Sect. 2. In such cases it is therefore necessary to apply other methods, based for example on the candidates' ages, draws and so on.

9 A Comparison with Other Methods

Tables 14 and 15 show the allocations assigned by the principal classical methods of rounding (presented in Sect. 3) in the cases of Examples 3 and 4, indicating some criteria which are violated.

Regarding Table 14, we add that all apportionments respect, at a local level, symmetry, monotonicity and Hare minimum; at global level Hare minimum and power index. Note that the new solutions respect all the criteria at global level; in particular the solutions in the last column respect all the criteria, contrary to classical methods.

Regarding Table 15, we add that all apportionments respect: at a local level, symmetry, monotonicity and power index; at a global level, monotonicity. Note that the new solutions respect all the criteria at a global level, contrary to classical methods.

10 Conclusions

The concept of solution proposed here avoids most of the distortions which arise when using classical methods and, when unavoidable, minimizes their negative effects. The procedures to obtain the solutions are simply applicable to automatic computation. The majority of classical techniques were developed before computers existed, or at least before they came into common use. We think it is now time to get up-to-date with electoral regulations, too.

Acknowledgments This work is sponsored by MIUR 40 % (PRIN 2005-prot. 2005149555_003). The authors thank Steven Brams, Francesc Carreras, Jacek Mercik and Hannu Nurmi for useful suggestions.

Appendix 1: Some Notes on the Banzhaf Normalized Power Index

In the Theory of Cooperative Games, a power index is a function which assigns shares of power to the players as a quantitative measure of their influence in voting situations.

For instance, suppose that a system is formed of three parties without particular propensity for special alliances, and that a simple majority is required. If the allotment of seats is (40, 30, 30), any reasonable power index will assign an equal power allotment of (1/3, 1/3, 1/3). If the seat allotment of the three parties is (60, 30, 10), then any reasonable index would give a power share of (1, 0, 0), since the first party attains the majority by itself. Some complications occur if the seat allotment is (50, 30, 20). If A, B, C are the three parties, we can remark that A is crucial for the three coalitions {A, B, C}, {A, B} and {A, C}, i.e. such coalitions attain the majority with party A and lose it without A. On the other hand, party B is only crucial for the coalition {A, B} and party C is only crucial for the coalition {A, C}. In general, the power indices are based on the crucialities of the parties. In particular, the Banzhaf index (1965) assigns to each party the number of coalitions for which it is crucial. In our example, the assigned powers are (3, 1, 1). The **Banzhaf normalized power index** assigns to each party a quota of the unity proportional to the number of coalitions for which it is crucial. In our example, the assigned powers are (3/5, 1/5, 1/5).

In addition to John F. Banzhaf, several authors independently introduced various indices having the same normalization: Coleman (1971), Penrose (1946) and, according to a particular interpretation, Luther Martin in the XVIII century [see Riker (1986) and Felsenthal and Machover (2005)]. That is the reason why this index should be mentioned as "Banzhaf-Coleman-Martin-Penrose Normalized power index".

A geometric interpretation is shown in Palestini (2005). For the automatic computation in general cases, we suggest the algorithm by Bilbao et al. (2000). The algorithm by Gambarelli (1996) takes into account previous computations, when the seats vary recursively. Then (with reference to the Appendix 2) it is more suitable for the application of the β_L-criterion, if computed before the N_L-criterion.

Overviews of further power indices can be found in Gambarelli (1983), Holubiec and Mercik (1994), Gambarelli and Owen (2004).

Appendix 2: An Algorithm Generator of the Solutions

We show an algorithm for the automatic generation of the solutions having as first criteria $F_G N_G \beta_G$ or $F_G \beta_G N_G$. Notice that this procedure can be easily structured for parallel processing, so that the time of computation can be considerably reduced.

INPUT

V the valid votes.

a the seats to be assigned to the districts.

"Global option" of the ordering of criteria at the global level ($F_G N_G \beta_G$ or $F_G \beta_G N_G$).

"Local option" of the ordering of criteria at the local level ($N_L \beta_L$ or $\beta_L N_L$).

OUTPUT

$S_1, S_2, \ldots, S_n$ the set of survived matrices.

WORKING AREAS

$\overline{N^V}$ the matrix of normalized votes.

$\overline{\beta^V}$ the matrix of the Banzhaf normalized power indices of votes.

$\overline{N^S}$ the matrix of normalized seats.

$\overline{\beta^S}$ the matrix of the Banzhaf normalized power indices of the seats.

$R_1, R_2, \ldots, R_n$ the set of matrices survived to the first local criterion.

B the set of vectors b generated by criteria $F_G N_G \beta_G$ or $F_G \beta_G N_G$.

c_{CUR} the vector $c(V, S)$ at the current step.

c_{MIN} the minimum vector c_{CUR} of the past steps.

f_d, f_p pointers to set S.

S the matrix in construction:

s_{11}	s_{12}	s_{13}	…	s_{1n_p}	a_1
s_{21}	s_{22}	s_{23}	…	s_{2n_p}	a_2
		…			…
$s_{n_d 1}$	$s_{n_d 2}$	$s_{n_d 3}$	…	$s_{n_d n_p}$	a_{n_d}
b_1	b_2	b_3	…	b_{n_p}	$\Sigma b_d = \Sigma a_p$

PROCEDURE
Read the input data.
Compute $\overline{V}$
Compute $\overline{\beta}$ using Bilbao et al. (2000).
Compute B according to the global option.
Set maximum values to c_{MIN}.

For every b of B:
    Set (n_d, n_p) as first pointers.
    Move n_d to f_d and n_p to f_p
    For all S of the current b:
    Set S (move a and b to the arrays of the totals and move zeroes to all s_{dp})
    Call the subroutine "Construction of the next S".
    Update f_d, f_p.
    Call the subroutine "Generation of solution" using R_k as output.
  Return
Return
Set maximum values to c_{MIN}.
Move n to m.
Varying t from 1 to m:
  Move R_t to S.
  Call subroutine "Generation of solution" using S_n as output.
Return
End

SUBROUTINE "GENERATION OF SOLUTION"

If the local option is N_L, compute $\overline{N^S}$

else compute $\overline{\beta^S}$ using Gambarelli (1996) (case $\beta_L N_L$) or Bilbao et al. (2000) (case $N_L \beta_L$).

During the above computation, construct c_{CUR} and compare it with c_{MIN}.
Just if $c_{\text{CUR}} > c_{\text{MIN}}$ exit.
When the construction of the normalized matrix is over:
If $c_{\text{CUR}} = c_{\text{MIN}}$ move $n + 1$ to n else move 1 to n move c_{CUR} to c_{MIN}.
Move S to output.
Exit

SUBROUTINE "CONSTRUCTION OF THE NEXT *S*"

If min $(a_{n_d}, b_{n_p}) = a_{n_d}$, then move a_{n_d} to $s_{n_d n_p}$,
move 0 to all the other elements of the last row and to a_{n_d},
move $(b_{n_p} - a_{n_d})$ to b_{n_p},
and iterate the procedure on the submatrix obtained by eliminating the last column, i.e. decreasing n_d by 1

If min $(a_{n_d}, b_{n_p}) = b_{n_p}$, then move b_{n_p} to $s_{n_d n_p}$,
move 0 to all other elements of the last column and to b_{np}
move $(a_{n_d} - b_{n_p})$ to a_{n_d},
and iterate the procedure on the submatrix obtained by eliminating the last column, i.e. decreasing n_p by 1

At the end of the procedure we obtain $a_1 = b_1$; this number will be moved to s_{11}
Exit

EXAMPLE

In example 3 the construction sequence of the first S is:

			6
			5
3	4	4	11

		0	6
		4	1
3	4	0	7

		0	6
0	1	4	0
3	3	0	6

	3	0	3
0	1	4	0
3	0	0	3

3	3	0	0
0	1	4	0
0	0	0	0

The sequence of the other S continues as follows:

2	4	0
1	0	4

3	2	1
0	2	3

3	1	2
0	3	2

...

0	2	4
3	2	0

References

Balinski, M. L., & Demange, G. (1989). An axiomatic approach to proportionality between matrices. *Mathematics of Operations Research, 14*, 700–719.

Balinski, M. L., & Young, H. P. (1982). *Fair Representation: Meeting the ideal of one man, one vote*. Yale University Press.

Banzhaf, J. F. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutger Law Review, 19*, 317–343.

Bilbao, J. M., Fernàndez, J. R., Jimenéz Losada, A., & Lopez, J. J. (2000). Generating functions for computing power indices efficiently. *Top, 8*(2), 191–213.

Brams, S. J. (1976). *Paradoxes in politics: An introduction to the nonobvious in political science*. New York: Free Press.

Coleman, J. S. (1971). Control of collectivities and the power of collectivities to act. In B. Lieberman (Ed.), *Social Choice* (Vol. 40, pp. 269–300). New York: Gordon and Breach.

Felsenthal, D. S., & Machover, M. (2005). Voting power measurement: A story of misreinvention. *Social Choice and Welfare, 25*(2–3), 485–506.

Gambarelli, G. (1983). Common behaviour of power indices. *International Journal of Game Theory, 12*, 237–244.

Gambarelli, G. (1996). Takeover algorithms. In M. Bertocchi, E. Cavalli & S. Komlosi (Eds.), Modelling techniques for financial markets and bank management: Proceedings of the 16th and 17th Euro Working Group of Financial Modelling Meetings, (pp. 212–222). Heidelberg: Physica Verlag.

Gambarelli, G. (1999). Minimax apportionments. *Group Decision and Negotiation, 8*, 441–461.

Gambarelli, G., & Owen, G. (2004). The coming of Game Theory. In G. Gambarelli (Ed.), *Essays on cooperative games: In honor of Guillermo Owen special issue of theory and decision* (Vol. 36, pp. 1–18). Dordrecht: Kluwer Academic Publishers.

Hodge, J. K., & Klima, R. E. (2005). *The Mathematics of Voting and Elections: A Hands-On Approach, Mathematical World 22*. American Mathematical Society.

Holubiec, J. W., & Mercik, J. W. (1994). Inside voting procedures. In *Studies in Economic and Social Sciences* (Vol. 2). Accedo Verlagsgesellschaft.

Nurmi, H. (1982). The problem of the right distribution of voting power. In M. J. Holler (Ed.), *Power, voting and voting power*. Würzburg: Physica-Verlag.

Palestini, A. (2005). Reformulation of some power indices in weighted voting games. *Power Measures II, Homo Oeconomicus, 22*, 487–507.

Penrose, L. S. (1946). The elementary statistics of majority voting. *Journal of the Royal Statistical Society, 109*, 53–57.

Riker, W. H. (1986). The first power index. *Social Choice and Welfare, 3*, 293–295.

Schmeidler, D. (1969). The nucleolus of a characteristic function game. *SIAM Journal of Applied Mathematics, 17*, 1163–1170.

Gridlock or Leadership in U.S. Electoral Politics

Evan Schnidman and Norman Schofield

1 Activist Politics

This chapter attempts to model elections by incorporating voter judgments about candidate and leader competence. In a sense the proposed model can be linked to Madison's understanding of the nature of the choice of Chief Magistrate (Madison 1999 [1787]) and Condorcet's work on the so-called "Jury Theorem" (Condorcet 1994 [1785]). This aspect of Condorcet's work has recently received renewed attention (McLennan 1998) and can be seen as a contribution to the development of a Madisonian conception of elections in representative democracies as methods of aggregation of both preferences and judgments.

The literature on electoral competition has focused on preferences rather than judgments. Models of two-party competition have typically been based on the assumption that parties or candidates adopt positions in order to win, and has inferred that parties will converge to the electoral *median*, under deterministic voting in one dimension (Downs 1957) or to the electoral mean in stochastic models.[1] This median based model has been applied recently by Acemoglu and Robinson (2000, 2006a) in a wide ranging account of political economy, including the transformation of the British polity to a more democratic model in the nineteenth century.

This chapter is an extension of Schofield and Schnidman (2011)

[1] See the earlier work by Enelow and Hinich (1989), Erikson and Romero (1990) and more recent work by Duggan (2006); Patty et al. (2009).

E. Schnidman (✉)
Department of Government, Harvard University, Cambridge, MA, USA
e-mail: evan.schnidman@gmail.com

N. Schofield
Center in Political Economy, Washington University, Saint Louis, MO, USA
e-mail: schofield.norman@gmail.com

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_11,

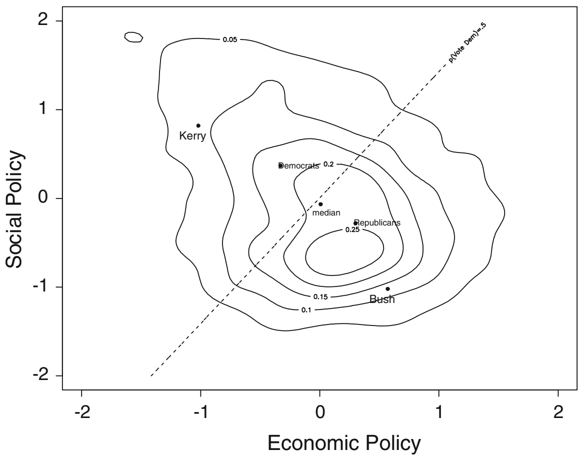


Fig. 1 Electoral distribution and candidate positions in the United States in 2004

In this chapter we develop a theory of political choice in which the political space is of higher dimension. This space is one which is derived essentially from the underlying factor structure of the political economy. That is to say, the axes are based on the preferences of those who control the factors of land, capital and labor. For example, Fig. 1 presents an estimate of the distribution of preferences (or preferred positions) in the U.S. presidential election of 2004.[2] The first-left right dimension represents preferences (or attitudes) towards government expenditure and taxes and can be interpreted as a *capital* axis. The second north-south or social dimension reflects attitudes on social policy, particularly civil rights, and can be interpreted as a *labor* axis.

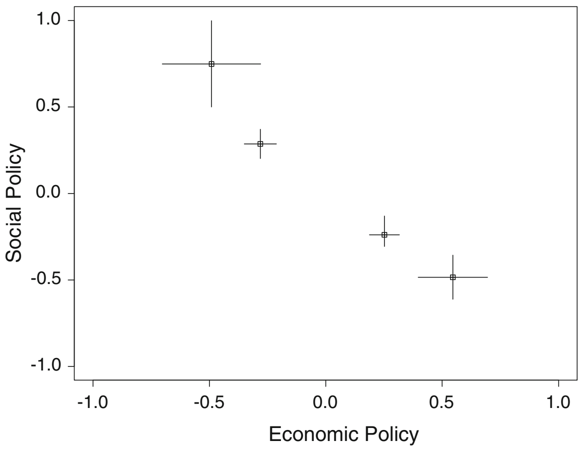
Because the political space is two-dimensional, parties in the United States must be coalitions of opposed interests. Figure 1 also shows a *partisan cleavage line* obtained from a simple logit model of the election. This cleavage line joins the preferred points of voters who, according to the stochastic vote model, would choose the candidates with equal probability of one half.

In Fig. 2 we present the results of a factor analysis of the 2004 ANES, showing estimated mean partisan and activist positions for Democrat and Republican voters in 2004 (error bars are larger for the mean activist positions. This Figure, together with Fig. 1 suggests that candidate positions are very much effected by activists who are estimated to be located at more extreme positions in the policy space. This inference is compatible with the model presented here.

The figure suggests that though the Republican party contains both socially conservative and socially liberal groups, they both tend to be *pro-capital*. Similarly the Democrat party tends to be *pro-labor*. The increasing dominance of "Tea Party" social conservatives in the Republic Party, and indeed the fact that the Republican position in the recent mid term election of 2010 appeared to be fairly

[2] This figure is based on factor analysis of the American National Election Study (ANES) for 2004 by Schofield et al. (2011a).

Fig. 2 Comparison of mean partisan and activist positions for Democrat and Republican voters in 2004 (error bars are larger for the mean activist positions



"radical" in the lower right quadrant of the political space, caused some prominent Republicans to consider a change of party allegiance to the Democrats. Shifts in the activist coalitions for the two parties thus cause a transformation of the *partisan cleavage line.*

Miller and Schofield (2003, 2008) argue that this is a fundamental aspect of U.S. politics: as activists on the "trailing edge"[3] of the cleavage line change party allegiance, then the positions of the two parties shift. This can be interpreted as a clockwise rotation in the political space. They suggest that in the 150 years since the Civil War, the partisan cleavage line has rotated nearly 180°, with the Republicans now occupying the position once occupied by the Democrats in the late nineteenth century. Miller and Schofield conjecture that in time, the Republican Party will adopt policies that are analogous to those proposed by William Jennings Bryan in 1896: populist and anti-business. In parallel, the Democratic Party will increasingly appeal to pro-business, social liberal cosmopolitans.

We argue that the fundamental changes in voter choice result not only from changes in the distribution of electoral preferences, but from the shifts in electoral *perceptions about the competence of the political candidates*.[4] These perceptions are influenced by the resources that the candidates command. In turn, these changes in perceptions are the consequence of the shifting pattern of activist support for the candidates.[5] The essence of the model presented here is that it attempts to endogenize the resources available to candidates by modeling the

---

[3] These would on the one hand be cosmopolitan, socially liberal but economically conservative Republicans or on the other hand, populist, socially conservative but economically liberal Democrats.

[4] In the empirical models that we have developed, perceptions are linked to candidate character traits such as moral, caring, knowledgable, honest, moral, strong, optimistic, intelligent.

[5] In recent elections, candidate resources are expended through the media. Even a hundred years ago, presidential candidates had to expend resources in campaigning throughout the country.

contracts they can make with their supporting activists. The activists must solve their own optimization problem by estimating the benefit they receive from their contributions and deciding what resources to make available to their chosen candidate.

In recent years, the importance of activist contributions has increased, and this has enhanced the influence of activist groups.[6] The empirical and formal models that we discuss here provide a reason why electoral politics has become so polarized in the United States.[7] Moreover, this polarization appears to have benefited the wealthy in society and may well account for the increase in inequality in income and wealth distribution that has occurred over the last decade (Hacker and Pierson 2006, 2010; Pierson and Skocpol 2007).

Essentially there is an arms race between candidates over these resources due to a feedback mechanism between politics and economics. As the outcome of the election becomes more important, activists become increasingly aware that the resources they provide have become crucial to election victories, and they become more demanding of their chosen candidates. Because of the offer of resources, candidates are forced to move to more radical positions, and polarization in candidate positions increases, even though there may be little change in the degree of polarization of the electorate.

Over the long run we see two forces at work: the continuing "circumferential" realignment and a "radial" polarization that occurs at times of political quandaries, caused by economic downturn or shocks to the global political economy.

In the next section we present an outline of the model that we use. In Sect. 3 we present the formal details of the model, and then in Sects. 4 and 5 we apply it to the consideration of the 2008 and 2010 elections in the United States. Section 6 applies the model to episodes in United States history, commenting on the balance between land, labor and capital. Section 7 concludes.

2 An Outline of the Model

In the standard spatial model, only candidate *positions* matter to voters. However, as Stokes (1963, 1992) has emphasized, the non-policy evaluations, or *valences*, of candidates by the electorate are equally important. In empirical models, a party's *valence* is usually assumed to be independent of the party's position, and adds to the statistical significance of the model. In general, valence reflects the overall degree to which the party is perceived to have shown itself able to govern

---

6 Indeed, Herrera et al. (2008) observe that spending by parties in federal campaigns went from 58 million dollars in 1976 to over 1 billion in 2004 in nominal terms.

7 See the works by Fiorina et al. (2005), Fiorina and Abrams (2009) and McCarty et al. (2006) on polarization in the electorate and Layman et al. (2010) on polarization among activists. Schofield et al. (2011a, b) gives similar results for the 2000, 2004 and 2008 elections.

effectively in the past, or is likely to be able to govern well in the future (Penn 2009).

Over the last decade a new literature has developed that considers deterministic or probabilistic voting models including valence or bias towards one or other of the candidates[8]

Recent work[9] has developed an empirical and formal stochastic electoral model based on multinomial conditional logit methods (MNL). In this model, each *agent*, j, was characterized by an *intrinsic or exogenous valence*, λ_j. This model can be considered to be Downsian, since it was based on a pure spatial model, where the estimates of valence were obtained from the intercepts of the model. It was possible to obtain the conditions for existence of "a local Nash equilibrium" (LNE) under vote maximization for a parallel formal model using the same stochastic assumptions as the MNL empirical model. A LNE is simply a vector of agent positions with the property that no agent may make a small unilateral move and yet increase utility (or vote share).

The *mean voter theorem* asserts that all candidates should converge to the electoral origin.[10] Empirical analyses of the 2000, 2004 and 2008 U.S. presidential elections (Schofield et al. 2011a, b) has corroborated the earlier work by Enelow and Hinich (1989) and shown, by simulation on the basis of the MNL models, that presidential candidates should converge to the electoral origin.[11] However, the empirical work also suggests that presidential candidates do not in fact adopt positions close to the electoral center. Figure 1, mentioned above, shows the estimated positions of the presidential candidates in the 2004 election in the U.S.

This figure is compatible with previous work empirical work by Poole and Rosenthal (1984) who also noted that there was no evidence of candidate convergence in U.S. presidential elections.

This chapter offers a more general model of elections that, we suggest, accounts for the difference between the estimates of equilibrium positions and actual candidate positions. The model is based on the assumption that there is a second kind of valence is known as *activist valence*. When party, or candidate j adopts a policy position z_j, in the policy space, X, then the activist valence of the party is denoted $\mu_j(z_j)$. Implicitly we adopt a model originally due to Aldrich (1983). In this model, activists provide crucial resources of time and money to their chosen party, and these resources are dependent on the party position.[12] The party then uses these resources to enhance its image before the electorate, thus affecting its overall

[8] Adams and Merrill (2005), Ansolabehere et al. (2001), Aragones and Palfrey (2002), Ashworth and Bueno de Mesquita (2009), Banks and Duggan (2005), Groseclose (2001) and McKelvey and Patty (2006).

[9] Schofield and Sened (2006), Schofield (2007).

[10] The electoral origin is the mean of the distribution of voter preferred points.

[11] Schofield et al. (2011c) obtains a similar result for the elections in Britain in 2005 and 2010.

[12] For convenience, it is assumed that $\mu_j(z_j)$ is only dependent on z_j, and not on $z_k, k \neq j$, but this is not a crucial assumption.

valence. Although activist valence is affected by party position, it does not operate in the usual way by influencing voter choice through the distance between a voter's preferred policy position, say x_i, and the party position. In this first model, as party j's activist support, $\mu_j(z_j)$, increases due to increased contributions to the party in contrast to the support $\mu_k(z_k)$ received by party k, then (in the model) all voters become more likely to support party j over party k.

The problem for each party is that activists are likely to be more extreme than the typical voter. By choosing a policy position to maximize activist support, the party will lose centrist voters. The party must therefore determine the "optimal marginal condition" to maximize vote share. The Theorem, presented in Sect. 3, gives this as a (first order) *balance condition*. Moreover, because activist support is denominated in terms of time and money, it is reasonable to suppose that the activist function will exhibit decreasing returns. The Theorem points out that when these activist functions are sufficiently concave, then the vote maximizing model will exhibit a Nash equilibrium.

It is intrinsic to the model that voters evaluate candidates not only in terms of the voters' preferences over intended policies, but also in terms of electoral judgements about the quality of the candidates. These judgements are in turn influenced by the resources that the candidates can raise from their activist supporters.

Grossman and Helpman (1996), in their game theoretic model of activists, consider two distinct motives for interest groups:

> Contributors with an *electoral motive* intend to promote the electoral prospects of preferred candidates, [while] those with an *influence motive* aim to influence the politicians' policy pronouncements.

In our first activist model the term $\mu_j(z_j)$ influences every voter and thus contributes to the electoral motive for candidate j. In addition, the candidate must choose a position to balance the electoral and activist support, and thus change the position adopted. This change provides the logic of activist influence.

We argue that the influence of activists on the two candidates can be characterized in terms of activist gradients.

Because each candidate is supported by multiple activists, we extend the activist model by considering a family of potential activists, $\{A_j\}$ for each candidate, j, where each $k \in A_j$ is endowed with a utility function, U_k, which depends on candidate j's position z_j, and the preferred position of the activist. The resources allocated to j by k are denoted $R_{jk}(U_k(z_j))$. Let $\mu_{jk}(R_{jk}(U_k(z_j)))$ denote the effect that activist k has on voters' utility. Note that the activist valence function for j is the same for all voters. With multiple activists, the *total activist valence function* for agent j is the linear combination $\mu_j(z_j) = \sum_{k \in A_j} \mu_{jk}(R_{jk}(U_k(z_j)))$. We also obtained information from the American National Election Surveys on activists-those who contributed resources to one or other of the two parties. Figure 1, above, showed the estimated positions of activists for the two parties. The figure does suggest that activists influence the candidate positions. The balance condition suggests that the aggregate activist gradients for each of the two candidates point into opposite quadrants of the policy space.

Bargains between the activists supporting candidate j then gives a *contract set* of activist support for candidate j, and this contract set can be used formally to determine the balance locus, or set of optimal positions for each candidate. This balance locus can then be used to analyze the pre-election contracts between each candidate and the family of activist support groups.

Consider now the situation where these contracts have been agreed, and each candidate is committed to a set of feasible contracts as outlined in Grossman and Helpman (1994, 1996, 2001). Suppose further that the activists have provided their resources. Then at the time of the election the effect of this support is incorporated into the empirical estimates of the various exogenous, sociodemographic and trait valences. Consequently, when we estimate these valences we also estimate the aggregate activist influence. The estimated positions of the candidates can then be regarded as incorporating policy preferences of the activists.

Electoral models where candidates have policy positions[13] implicitly assume that candidates would be willing to accept defeat because of an adherence to particular policy positions. We argue that it is more plausible that the estimated positions of the candidates are the result of maximizing candidate utility functions that balance the electoral consequences of position-taking with the necessity of obtaining activist resources to contest the election. This calculation requires an estimate of the degree to which these resources will influence the perceptions that the electorate has of the various valences associated with the model.

In the final version of the model we allow the activist valence function to be individual specific. The total resources available to candidate j are now denoted $\boldsymbol{\mu}_j(z_j)$, and these may be allocated to individuals, with resource m_{ij} targeted on voter, or "voter class", i by candidate j. Since m_{ij} will depend on z_j, we write this allocation as $m_{ij}(z_j)$, so the budget constraint is

$$\begin{aligned}\mathbf{R}_j(z_j) &= \sum_{k\in A_j} R_{jk}(U_k(z_j)) \\ &= \sum_{i\in N} m_{ij}(z_j).\end{aligned}$$

The optimization problem is now a more complex one, subject to this constraint. In actual fact, candidates will generally not allocate resources to individuals per se, but to voter classes via media outlets in different regions, or "zip codes." The general balance condition specifies how these resources should be allocated throughout the polity.

[13] As proposed by Wittman (1977), Calvert (1985), Duggan and Fey (2005), Duggan (2006) and Peress (2010).

A recent literature on elections has focussed on the effects of campaign expenditure on US election results (Coate 2004).[14] Herrera et al. (2008) suggest that electoral volatility forces candidates to spend more, while Ashworth and Bueno de Mesquita (2009) suppose that candidates buy valence so as to increase their election chances. Meirowitz (2008) notes that "candidates and parties spending this money thought that it would influence the election outcome. Downsian models of competition cannot explain how candidates choose campaign spending levels or what factors influence these decisions." Meirowitz proxies the choice of expenditure in terms of candidate choice of effort, but his model does not explicitly deal with an endogenous budget constraint.

3 The Formal Stochastic Model

3.1 The First Activist Model

We develop an electoral model that is an extension of the multiparty stochastic model of McKelvey and Patty (2006), modified by inducing asymmetries in terms of valence. The justification for developing the model in this way is the empirical evidence that valence is a natural way to model the judgements made by voters of party leaders and candidates. There are a number of possible choices for the appropriate model for multiparty competition. The simplest one, which is used here, is that the utility function for the candidate of party j is proportional to the vote share, V_j, of the party in the election.[15] With this assumption, we can examine the conditions on the parameters of the stochastic model which are necessary for the existence of a pure strategy Nash equilibrium (PNE). Because the vote share functions are differentiable, we use calculus techniques to obtain conditions for positions to be locally optimal. Thus we examine what we call *local pure strategy Nash equilibria* (LNE). From the definitions of these equilibria it follows that a PNE must be a LNE, but not conversely.

The key idea underlying the formal model is that party leaders attempt to estimate the electoral effects of policy choices, and choose their own positions as best responses to other party declarations, in order to maximize their own vote share. The stochastic model essentially assumes that candidates cannot predict vote response precisely, but that they can estimate the effect of policy proposals on the expected vote share. In the model with valence, the stochastic element is

---

[14] An earlier chapter by Groseclose and Snyder (1996) looked at vote buying, but in the legislature.

[15] For refining the model, and for empirical analysis, it would be more appropriate to use the share of the electoral college votes, or a combination of this and the party vote shares in the elections to Congress. We adopt this simplifying assumption in order to present the essential structure of the formal model.

associated with the weight given by each voter, i, to the average perceived quality or valence of the candidate.

Definition 1 The Stochastic Vote Model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\alpha}, \boldsymbol{\mu}, \beta; \Psi)$ with Activist Valence.

The data of the spatial model is a distribution, $\{x_i \in X\}_{i \in N}$, of voter ideal points for the members of the electorate, N, of size n. We assume that X is a compact convex subset of Euclidean space, $\mathbb{R}^w$, with w finite. Without loss of generality, we adopt coordinate axes so that $\frac{1}{n}\Sigma x_i = 0$. By assumption $0 \in X$, and this point is termed the *electoral mean*, or alternatively, the *electoral origin*. Each of the parties in the set $P = \{1, \ldots, j, \ldots, p\}$ chooses a policy, $z_j \in X$, to declare prior to the specific election to be modeled.

Let $\mathbf{z} = (z_1, \ldots, z_p) \in X^p$ be a typical vector of candidate policy positions.

We define a stochastic electoral model, which utilizes socio-demographic variables and voter perceptions of character traits. For this model we assume that voter i utility is given by the expression

$\mathbf{u}_i(x_i, \mathbf{z}) = (u_{i1}(x_i, z_1), \ldots, u_{ip}(x_i, z_p))$ where

$$u_{ij}(x_i, z_j) = \lambda_j + \mu_j(z_j) + (\theta_j \cdot \eta_i) + (\alpha_j \cdot \tau_i) - \beta \|x_i - z_j\|^2 + \epsilon_j \tag{1}$$

$$= u^*_{ij}(x_i, z_j) + \epsilon_j. \tag{2}$$

Here $u^*_{ij}(x_i, z_j)$ is the observable component of utility. The constant term, λ_j, is the *intrinsic or exogenous valence* of party j, The function $\mu_j(z_j)$ is the component of valence generated by activist contributions to agent j. The term β is a positive constant, called the *spatial parameter*, giving the importance of policy difference defined in terms of a metric induced from the Euclidean norm, $|| \cdot ||$, on X. The vector $\epsilon = (\epsilon_1, \ldots, \epsilon_j, \ldots, \epsilon_p)$ is the stochastic error, whose multivariate cumulative distribution is the Type 1 extreme value distribution, denoted by Ψ.

Sociodemographic aspects of voting are modeled by $\boldsymbol{\theta}$, a set of k-vectors $\{\theta_j : j \in P\}$ representing the effect of the k different sociodemographic parameters (class, domicile, education, income, religious orientation, etc.) on voting for party j while η_i is a k-vector denoting the i th individual's relevant "sociodemographic" characteristics. The compositions $\{(\theta_j \cdot \eta_i)\}$ are scalar products, called the *sociodemographic valences* for j.

The terms $(\alpha_j \cdot \tau_i)$ are scalars giving voter i's perception of the *traits* of the leader (or candidate) of party j. The coefficients, α_j, correspond to different candidates. We let $\boldsymbol{\alpha} = (\alpha_p, \ldots \alpha_1)$.[16] The trait score can be obtained by factor analysis from a set of survey questions asking respondents about the traits of the candidate. including moral, caring, knowledgable, strong, dishonest, intelligent, out of touch. Schofield et al. (2011a, b) show that the electoral perceptions of candidate traits are statistically relevant for modeling US presidential elections.

[16] For US elections we talk of the traits of candidate j, rather than party leader j.

It is assumed that the intrinsic valence vector

$$\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_p) \text{ satisfies } \lambda_p \geq \lambda_{p-1} \geq \cdots \geq \lambda_2 \geq \lambda_1.$$

Voter behavior is modeled by a probability vector. The probability that a voter i chooses party j at the vector $\mathbf{z}$ is

$$\rho_{ij}(\mathbf{z}) = \Pr[[u_{ij}(x_i, z_j) > u_{il}(x_i, z_l)], \text{ for all } l \neq j]. \tag{3}$$

$$= \Pr[\epsilon_l - \epsilon_j < u^*_{ij}(x_i, z_j) - u^*_{il}(x_i, z_l), \text{ for all } l \neq j]. \tag{4}$$

Here Pr stands for the probability operator generated by the distribution assumption on ϵ. The *expected vote share* of agent j is

$$V_j(\mathbf{z}) = \frac{1}{n} \sum_{i \in N} \rho_{ij}(\mathbf{z}). \tag{5}$$

The differentiable function $V : X^p \to \mathbb{R}^p$ is called the *party profile function*.

The most common assumption in empirical analyses is that Ψ is the *Type I extreme value distribution* (also called the Gumbel (maximum) distribution). The theorem in this chapter is based on this assumption. This distribution assumption is the basis for much empirical work based on multinomial logit estimation.

Definition 2 The Type I Extreme Value Distribution, Ψ.

(i) The cumulative distribution, Ψ, has the closed form

$$\Psi(h) = \exp[-\exp[-h]],$$

with probability density function

$$\psi(h) = \exp[-h] \exp[-\exp[-h]]$$

and variance $\frac{1}{6}\pi^2$.

(ii) For each voter i, and party j, the probability that a voter i chooses party j at the vector $\mathbf{z}$ is

$$\rho_{ij}(\mathbf{z}) = \frac{\exp[u^*_{ij}(x_i, z_j)]}{\sum_{k=1}^{p} \exp u^*_{ik}(x_i, z_k)}. \tag{6}$$

See Train (2003 :79). In this stochastic electoral model it is assumed that each party j chooses z_j to maximize V_j, conditional on $\mathbf{z}_{-j} = (z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_p)$.

Definition 3 Equilibrium Concepts.

(i) A vector $\mathbf{z}^* = (z_1^*, \ldots, z_{j-1}^*, z_j^*, z_{j+1}^*, \ldots, z_p^*)$ is a *local Nash equilibrium* (LNE) if, for each agent j, there exists a neighborhood X_j of z_j^* in X such that

$$V_j(z_1^*, \ldots, z_{j-1}^*, z_j^*, z_{j+1}^*, \ldots, z_p^*) \geq V_j(z_1^*, \ldots, z_j, \ldots, z_p^*) \text{ for all } z_j \in X_j.$$

(ii) A vector $\mathbf{z}^* = (z_1^*, \ldots, z_{j-1}^*, z_j^*, z_{j+1}^*, \ldots, z_p^*)$ is a *pure strategy Nash equilibrium* (PNE) iff X_j can be replaced by X in (i)..

(iii) The strategy z_j^* is termed a *local strict best response*, a *local weak best response*, or a *global best response*, respectively to $\mathbf{z}_{-j}^* = (z_1^*, \ldots, z_{j-1}^*, z_{j+1}^*, \ldots, z_p^*)$ depending on which of the above conditions is satisfied.

(iv) Strict local Nash equilibria (SLNE) and strict Nash equilibria (SPNE) are defined analogously by requiring strict inequalities in the definition.

From the definitions, it follows that if $\mathbf{z}^*$ is a PNE it must be an LNE.

Notice that in this model, each agent is uncertain about the precise electoral outcome, because of the stochastic component of voter choice. None the less, we presume that each agent uses opinion poll data, etc. to estimate expected vote share, and then responds to this information by searching for a “local equilibrium” policy position in order to gain as many votes as possible.

It follows from (6) that for voter i, with ideal point, x_i, the probability, $\rho_{ij}(\mathbf{z})$, that i picks j at $\mathbf{z}$ is given by

$$\rho_{ij}(\mathbf{z}) = \left[1 + \Sigma_{k \neq j}[\exp(f_{jk})]\right]^{-1} \tag{7}$$

$$\text{where } f_{jk} = u_{ik}^*(x_i, z_j) - u_{ij}^*(x_i, z_j).$$

We use (9) to show that the first order condition for $\mathbf{z}^*$ to be a LNE is that it be a *balance solution*.

Definition 4 The balance solution for the model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\alpha}, \boldsymbol{\mu}, \beta; \Psi)$.

Let $[\rho_{ij}(\mathbf{z})] = [\rho_{ij}]$ be the n by p matrix of voter probabilities at the vector $\mathbf{z}$, and let

$$[\varpi_{ij}] = \left[\frac{\rho_{ij} - \rho_{ij}^2}{\Sigma_{k=1}^n(\rho_{kj} - \rho_{kj}^2)}\right] \tag{8}$$

be the n by p matrix of weighting coefficients.

The *balance equation* for z_j^* is given by expression

$$z_j^* = \frac{1}{2\beta}\frac{d\mu_j}{dz_j} + \sum_{i=1}^n \varpi_{ij} x_i. \tag{9}$$

The vector $\sum_i \varpi_{ij} x_i$ is a convex combination of the set of voter ideal points. This vector is called the *weighted electoral mean* for party j. Define

$$z_j^{el} = \sum_i \varpi_{ij} x_i. \tag{10}$$

The balance equations for $j = 1, \ldots, p$ can then be written as

$$\left[z_j^{el} - z_j^* \right] + \frac{1}{2\beta} \frac{d\mu_j}{dz_j} = 0. \tag{11}$$

The bracketed term on the left of this expression is termed the *marginal electoral pull of party j* and is a gradient vector pointing from z_j^* towards the *weighted electoral mean*, z_j^{el}, of the party. This weighted electoral mean is that point where the electoral pull is zero. Notice that the each entry in the vector $\mathbf{z}^{el} = (z_1^{el}, z_2^{el}, \ldots z_p^{el})$ depends on all other entries. The vector $\frac{d\mu_j}{dz_j}$ is called *the marginal activist pull for party j*.

If $\mathbf{z}^*$ satisfies the system of balance equations, for all j, then call $\mathbf{z}^*$ a *balance solution*.

For the following discussion note again that by suitable choice of coordinates, the equi-weighted electoral mean $\frac{1}{n}\Sigma x_i = 0$, and is termed the *electoral origin*.

The following theorem is proved in Schofield (2006a) .

Activist Theorem 1

Consider the electoral model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\alpha}, \boldsymbol{\mu}, \beta; \Psi)$ based on the Type I extreme value distribution, and including both intrinsic and activist valences.

(i) The first order condition for $\mathbf{z}^*$ to be an LNE is that it is a balance solution.
(ii) If all activist valence functions are highly concave, in the sense of having negative eigenvalues of sufficiently great magnitude, then a balance solution will be a LNE.

Notice that if X is open, then this first order condition at $\mathbf{z}^*$ is necessary for $\mathbf{z}^*$ to be a PNE.

3.2 Extension to the Case with Multiple Activist Groups

(i) For each party leader, j, let $\{A_j\}$ be a family of potential activists, where each $k \in A_j$ is endowed with a utility function, U_k, which is a function of the position z_j. The resources allocated to j by k are denoted $R_{jk}(U_k(z_j))$. The total activist valence function for leader j is the linear combination

$$\mu_j(z_j) = \sum_{k \in A_j} \mu_{jk}(R_{jk}(U_k(z_j))). \tag{12}$$

where $\{\mu_{jk}\}$ are functions of the contributions $\{R_{jk}(U_k(z_j))\}$, and each μ_{jk} is a concave function of R_{jk}.

(ii) Assume the gradients of the valence functions for j are given by

$$\frac{d\mu_{jk}}{dz_j} = a_k^* \frac{dR_{jk}}{dz_j} = a_k^* a_k^{**} \frac{dU_k}{dz_j} \tag{13}$$

where the coefficients, $\{a_k^*, a_k^{**}\} > 0$, and are differentiable functions of z_j.

(iii) Under these assumptions, the first order equation $\frac{d\mu_j}{dz_j} = 0$ becomes

$$\frac{d\mu_j}{dz_j} = \sum_{k \in A_j} \frac{d}{dz_j}[\mu_{jk}(R_{jk}(U_k(z_j)))] \tag{14}$$

$$= \sum_{k \in A_j} (a_k^{**} a_k^*) \frac{dU_k}{dz_j} = 0. \tag{15}$$

The *Contract Curve* generated by the family $\{A_j\}$ is the locus of points satisfying the gradient equation

$$\sum_{k \in A_j} b_k \frac{dU_k}{dz_j} = 0, \text{ where } \sum_{k \in A_j} b_k = 1 \text{ and all } a_k > 0. \tag{16}$$

Here we let $b_k = (a_k^{**} a_k^*)$ and renormalize.

The *Balance Locus* for the leader j, defined by the family, $\{A_j\}$, is the solution to the first-order gradient equation

$$\left[z_j^{el} - z_j^*\right] + \frac{1}{2\beta}\left[\sum_{k \in A_j} a_k \frac{dU_k}{dz_j}\right] = 0. \tag{17}$$

The simplest case, discussed in Miller and Schofield (2003) is in two dimensions, where each leader has two activist groups.[17] In this case, the contract curve for each leader's supporters will, generically, be a one-dimensional arc. Miller and Schofield also supposed that the activist utility functions were ellipsoidal, mirroring differing saliences on the two axes. In this case the contract curves would be *catenaries*, and the balance locus would be a one dimensional arc. The balance solution for each leader naturally depends on the positions of opposed leaders, and

[17] See also Schofield and Cataife (2007) for example.

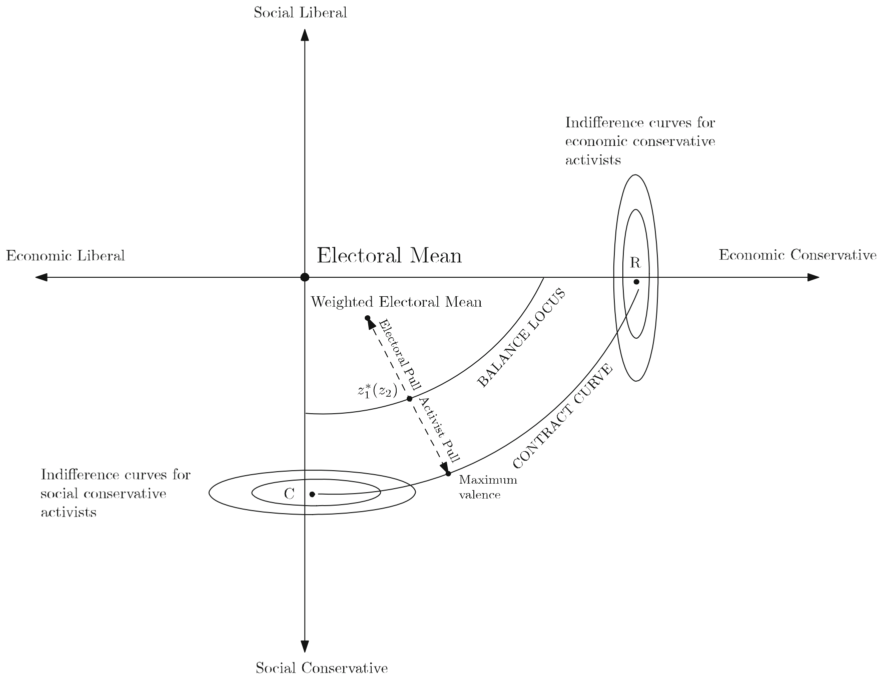


Fig. 3 Optimal Republican position

on the coefficients, as indicated above, of the various activists. The determination of the balance solution can be obtained by computing the vote share Hessian along the balance locus. Because the activist valence functions can be expected to be concave in the activist resources, the Hessian of the overall activist valence, μ_j, can be expected to have negative eigenvalues. For this reason, the Activist Theorem 1 gives a formal reason to expect existence of a PNE. In Fig. 3, the point $z_1^*(z_2)$ satisfies the balance equation for a Republican candidate. This point lies on the balance locus of the Republican party, and is also a function of the Democrat candidate location, z_2. A similar balance locus can be constructed for the Democrat candidate. Note that Fig. 1 is compatible with Fig. 3.

If we associate the utilities $\{U_k\}$ with leaders of the activist groups for the parties, then the combination

$$\sum_{k \in A_j} a_k \frac{dU_k}{dz_j}$$

may be interpreted as the marginal utility of the candidate for party j, induced by the activist support. Notice that the model presented here is formally identical to one where the party leader has policy preferences. This activist model can be given a game-theoretic foundation, as in Grossman and Helpman (2001), and can in principle be extended to the case where there are multiple activist groups which have the option of choosing from among a set of possible party leaders, all with varying intrinsic valences and preferences (Galiani et al. 2012).

3.3 Extension of the Activist Model: Targeting Voters

We now reinterpret

$$\boldsymbol{\mu}_j(z_j) = \sum_{k \in A_j} R_{jk}(U_k(z_j)). \tag{18}$$

as the total resources obtained by agent j from the various activist groups. These resources are denominated in terms of time (times skilled labor rate) or money, so we can take the units as monetary.

These resources are used to target the individual voters and the voter utility function is now

$$\begin{aligned} u_{ij}(x_i, z_j) &= \lambda_j + \mu_i(m_{ij}) + (\theta_j \cdot \eta_i) + (\alpha_j \cdot \tau_i) - \beta \|x_i - z_j\|^2 + \varepsilon_j \\ &= u^*_{ij}(x_i, z_j) + \varepsilon_j. \end{aligned}$$

Here $\mu_i(m_{ij})$ is the valence effect of the expenditure of resources, (m_{ij}) on the targeting of voter i, by agent j. We assume that the greater the resources m_{ij} spent on persuading voter i, the greater the implicit valence associated with candidate j, so $\frac{d\mu_i(m_{ij})}{dm_j} > 0$. We may also assume decreasing returns so that $\frac{d^2\mu_i(m_{ij})}{dm_j^2} < 0$. Obviously we can partition the voters into different categories, in terms of their sociodemographic valences. Note that different agents may target the same voter or group of voters.

We assume that for each j the budget constraint is satisfied:

$$\mathbf{R}_j(z_j) = \sum_{k \in A_j} R_{jk}(U_k(z_j)) = \sum_{i \in N} m_{ij} \tag{19}$$

We now assume that j solves the optimization problem that we now construct. Since $\boldsymbol{\mu}_j(z_j)$ determines the budget constraint for j, we can write $m_{ij} \equiv m_{ij}(z_j)$, so

$$\mu_i(m_{ij}) \equiv \mu_i(m_{ij}(z_j) \equiv \mu_{ij}(z_j).$$

We shall also assume that the solution to the optimization problem is smooth, in the sense that $\mu_{ij}()$ is a differentiable function of z_j.

Then just as above, the first order condition gives a more general balance condition as follows:

$$0=\frac{dV_j(\mathbf{z})}{dz_j}=\frac{1}{n}\sum_{i\in N}\frac{d\rho_{ij}}{dz_j}$$
$$=\frac{1}{n}\sum_{i\in N}[\rho_{ij}-\rho_{ij}^2]\left\{2\beta(x_i-z_j)+\frac{d\mu_{ij}}{dz_j}(z_j)\right\}.$$

$$\text{So}\, z_j\sum_{i\in N}[\rho_{ij}-\rho_{ij}^2]=\sum_{i\in N}[\rho_{ij}-\rho_{ij}^2]\left\{x_i+\frac{1}{2\beta}\frac{d\mu_{ij}}{dz_j}(z_j)\right\}.$$

$$\text{Hence}\, z_j^*=\frac{\sum_i\left[[\rho_{ij}-\rho_{ij}^2]\left[x_i+\frac{1}{2\beta}\frac{d\mu_{ij}}{dz_j}(z_j)\right]\right]}{\sum_{k\in N}[\rho_{kj}-\rho_{kj}^2]}$$
$$\text{or}\, z_j^*=\sum_{i=1}^n\varpi_{ij}(x_i+\gamma_i)\text{ where }\gamma_i=\frac{1}{2\beta}\frac{d\mu_{ij}}{dz_j}(z_j)$$

$$\text{and}\,\varpi_{ij}=\frac{[\rho_{ij}-\rho_{ij}^2]}{\sum_{k\in N}[\rho_{kj}-\rho_{kj}^2]}$$

This can be written $\left[z_j^*-z_j^{el}\right]=\sum_{i=1}^n\varpi_{ij}\gamma_i$ where $z_j^{el}=\sum_{i=1}^n\varpi_{ij}x_i$.

$$\text{When}\,\frac{d\mu_{ij}}{dz_j}(z_j)=\frac{d\mu_j}{dz_j}(z_j)$$

this reduces to the previous result (11).

The difference now is that instead of there being a single *centrifugal marginal activist pull* $\frac{1}{2\beta}\frac{d\mu_j}{dz_j}(z_j)$ *there is an aggregate activist pull*

$$\sum_{i=1}^n\varpi_{ij}\gamma_i=\frac{1}{2\beta}\sum_{i=1}^n\frac{[\rho_{ij}-\rho_{ij}^2]}{\sum_{k\in N}[\rho_{kj}-\rho_{kj}^2]}\frac{d\mu_{ij}}{dz_j}(z_j)$$

determined by the budget constraint given in Eq. (13).

Notice that the first order condition depends on the marginal terms, $\frac{d\mu_{ij}}{dz_j}(z_j)$, associated with policy positions, and these will depend on the marginal valence effects $\frac{d\mu_i(m_{ij})}{dm_j}$. Although these valence effects can be assumed to exhibit decreasing returns, these will vary across different classes of voters. The plausibility of existence of Nash equilibria turns on whether the induced second order terms

$\frac{d^2\mu_{ij}}{dz_j^2}(z_j)$ have negative eigenvalues. The assumption of negative eigenvalues would give a version of the activist theorem.

Note also that if ρ_{ij} is close to 0 or 1, then ϖ_{ij} will be close to 0, so the optimal calculation will be complex, though in principle solvable. It is plausible the candidate should expend resources on pivotal voters for whom ρ_{ij} is close to 1/2.[18]

To sketch an outline of a general model to endogenize activist support, we first let

$$\boldsymbol{\rho} : [X \times \mathbb{B}^n]^p \to [0, 1]^{n \times p}$$

specify the voter probabilities in terms of candidate positions in X^p and the distribution, in $\mathbb{B}^{n \times p}$, of resources $\{m_{ij}\}$ to all voters.[19]

We then let

$$\mathbf{V} = V_1 \times .. \times V_p : [X \times \mathbb{B}^n]^p \to [0, 1]^p$$

be the *party profile function*, mapping party positions and voter distributions to vote shares, as given by the above models. Indeed, for a more general model we could consider multiparty systems where agents form beliefs about coalitions behavior, as suggested in Schofield and Sened (2006). In this case the mapping would be

$$\mathbf{V} = V_1 \times .. \times V_p : [X \times \mathbb{B}^n]^p \to \mathbb{R}^p.$$

We assume that each of the **k** activists offers a distribution of resources to the p party leaders, which we take to be a vector in $\mathbb{B}^{\mathbf{k}}$. We seek is an equilibrium to a game form which may be written

$$\mathbf{U} \otimes \mathbf{V} : \mathbf{W} = B^{\mathbf{k}} \times [X \times \mathbb{B}^n]^p \to \mathbb{R}^{\mathbf{k}} \times \mathbb{R}^p.$$

This is an extremely complex dynamical game, and we do not attempt to explore the full ramifications of this model here.[20] One way to deal with it is to consider a dynamical version by considering a preference field for each party, or activist. This will be a cone in the tangent space of the agent's strategy space which specifies those changes in the agent's behavior which increase the agents utility. We denote the joint preference field by

---

[18] Stokes (2005) make a somewhat similar inference, discussing clientist models of politics, where m_{ij} is simply a monetary bribe to i. Obviously the marginal benefit to a poor voter is greater than to a wealthy voter, under the usual assumption of decreasing marginal utility for money.

[19] It is reasonable to assume that the resource distributions lie in a compact ball, namely $\mathbb{B}^{n \times p}$.

[20] See Coram (2010) for a dynamical version of a similar model. Acemoglu and Robinson (2008) also develop a model based on Markov Perfect Equilibrium where the elite, activists, have different preferences for the public good, in X and contribute to the *de facto* power of the political leader. However, they do not assume competing political leaders. The "matching" model proposed by Jackson and Watts (2010) embeds the Nash equilibrium within a coalition game, and would allow the principals to switch from one agent coalition to another.

$$H_{\mathbf{U}\otimes\mathbf{V}} : \mathbf{W} \to \mathbb{T}\mathbf{W}$$

where $\mathbb{T}\mathbf{W}$ stands for the tangent bundle above $\mathbf{W}$. A result in Schofield (2011) shows that if the tangent field $H_{\mathbf{U}\otimes\mathbf{V}}$ satisfies a "half open property" then there will exist a critical Nash equilibrium satisfying the first order condition for equilibrium.

Earlier results of Schofield (1978), McKelvey (1979) had suggested chaos could be generic in electoral models.[21] The application of this model (in Sect. 6) to the historical development of the U.S. political economy suggests that the equilibria of the model are subject to both circumferential and radial transformations over time.

4 Activist Support for Parties in the United States

To apply the above model, suppose there are two dimensions of policy, one economic, and one social. These can be found by factor analysis of survey data.

As in Fig. 2 indicates, we represent the conflicting interests or bargains between the two activist groups of supporters for the Republican Party, located at R and C, by a "contract curve." This represents the set of policies that these two groups would prefer their candidate to adopt. It can be shown that this contract curve is a *catenary* whose curvature is determined by the eccentricity of the utility functions of the activist groups (Miller and Schofield 2003). We call this the *Republican contract curve*. The Democrat activist groups may be described by a similar contract curve (This is the simplest case with just two activist groups for each candidate. As the previous section shows, this idea can be generalized to many activist groups.)

The theorem presented above gives the first order condition for the candidate positions (z^*_{dem}, z^*_{rep}) to be a Nash equilibrium in the vote share maximizing game. This condition is that the party positions satisfy a *balance equation*. This means that, for each party, $j = dem$ or rep, there is a weighted electoral mean for party j, given by the expression

$$z^{el}_j = \sum_i \varpi_{ij} x_i. \tag{20}$$

This is determined by the set of voter preferred points $\{x_i\}$. Notice that the coefficients $\{\varpi_{ij}\}$ for candidate j will depend on the position of the other candidate, k. As presented in the formal model, the *balance equation* for each j is given by:

$$\left[z^{el}_j - z^*_j\right] + \frac{1}{2\beta}\left[\frac{d\mu_j}{dz_j}|_z\right] = 0. \tag{21}$$

[21] See also Riker (1980, 1982, 1986).

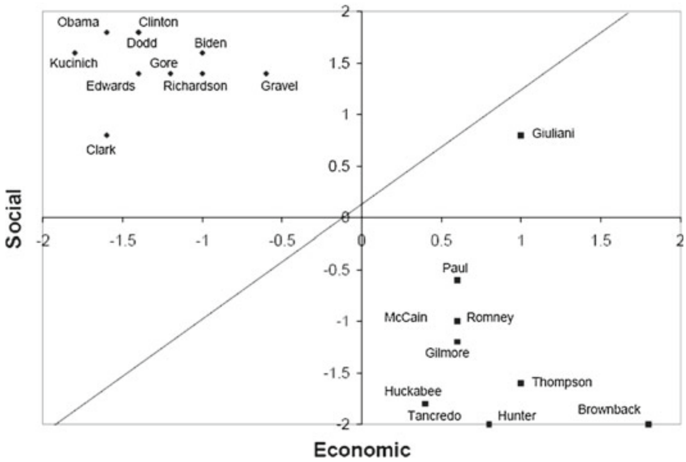


Fig. 4 Positions of Republican and Democrat candidates in 2008

4.1 The 2008 Election

The previous sections have suggested that a candidate's valence at election time is due to the ability of activist groups to raise resources for the candidate. At the same time, the candidate positions are the result of a balancing act between choosing an electorally optimal position and being able to persuade activist groups to provide these resources. We briefly provide some information about this balancing act: Fig. 4 shows the estimated positions of Republican and Democrat Presidential primary candidate positions prior to the 2008 election. The figure clearly suggests that Obama adopted a fairly extreme policy position, very liberal on both economic and social axes. Figures 5 and 6 show the relationship between electoral popularity of the candidates and their campaign expenditures as of January 2008.

Figure 5 suggests that Obama and Hilary Clinton were both very successful in raising campaign resources, and that these were highly correlated with the electoral support. Other candidates fell far behind and dropped out of the race. Figure 6 suggests that McCain was also extremely popular, even though his campaign, in January 2008, had not been very successful in raising contributions. This inference is compatible with McCain's estimated fairly moderate position in Fig. 4. Obviously, the relationship between campaign resources and popular vote in primaries and in the general election is extremely complex. Further research will attempt to utilize the model presented here to clarify this relationship.

Obama's victory on November 4, 2008 suggests that it was the result of an overall shift in the relative valences of the Democrat and Republican candidates from the election of 2004. Indeed, Schofield et al. (2011b) analyse a spatial model of the 2008 election and obtain a figure of 0.84 for the estimate of Obama's valence advantage over McCain.

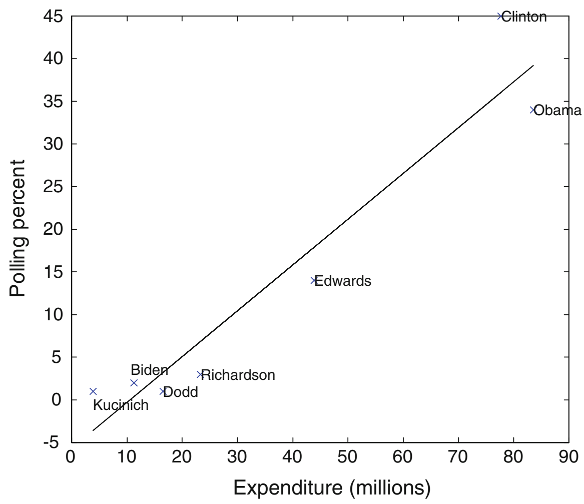


Fig. 5 Democrat candidate spending and popularity, January 2008

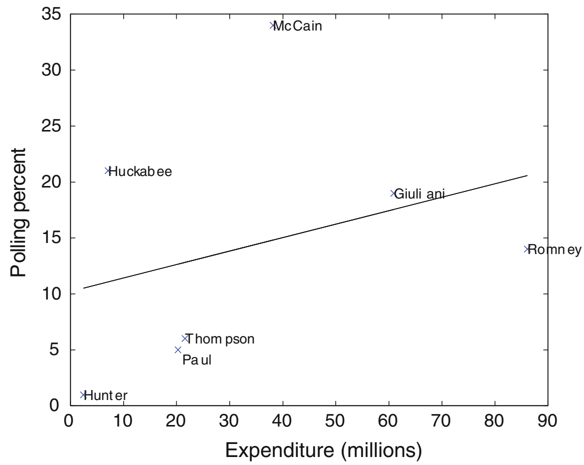


Fig. 6 Republican candidate spending and popularity, January 2008

In fact there were differential shifts in different regions of the country. In a region of the country from West Virginia through Tennessee, Arkansas and Oklahoma, there was an increase of 20 % in the Republican vote over the share for 2004, suggesting a regional change of about 0.6 in McCain's valence advantage.

Obama's victory in 2008 suggests that policy outcomes during his administration ought to lie in the upper left hand quadrant of the policy space. Figure 7 provides an estimate (taken from Schofield et al. 2011a) of the location of McCain and Obama at the November 2008 election. The Figure also shows the location of Democrat and Republican activists. Again, there is some evidence that extreme activists influence the policy choices of the candidates.

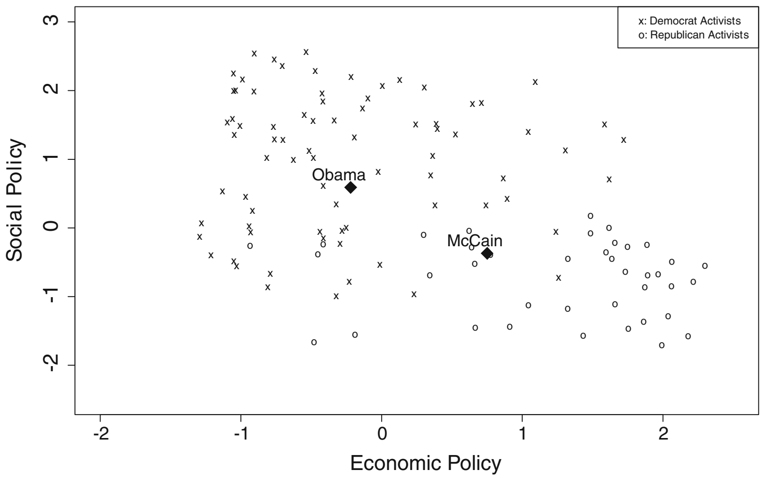


Fig. 7 Estimated US Presidential candidate positions in 2008 and activist positions

5 Post 2008 Election

The precise policy outcome from Obama's administration have thus far depended largely on the degree to which Republicans in the Senate have blocked Democratic policies through the use of the filibuster. Early in his administration some of Obama's policy initiatives successfully passed through Congress but only after navigating Republican opposition in the Senate. For example, on January 15, 2009, the Senate voted 52 against and 42 in support of Obama's economic recovery program. On February 6, 2009 an agreement was reached in the Senate to reduce the size of the stimulus bill to $780 billion, in return for the support of three Republican senators. On February 9 the Senate did indeed vote by the required majority of 61 to halt discussion of the stimulus bill, thus blocking a filibuster. A compromise bill of $787 billion, including some tax cuts, was agreed upon by both the House and Senate within a few days; the bill passed the House with 245 Democrats voting in favor and 183 Republicans voting against while the Senate passed it with just 60 votes. The bill was immediately signed by President Obama.

As Obama commented afterwards:

> Now I have to say that given that [the Republicans] were running the show for a pretty long time prior to me getting there, and that their theory was tested pretty thoroughly and its landed us in the situation where we've got over a trillion dollars' worth of debt and the biggest economic crisis since the Great Depression, I think I have a better argument in terms of economic thinking.

On February 26, 2009 Obama proposed a 10 year budget that revised the priorities of the past, with an estimated budget deficit for 2009 at $1.75 trillion (over 12 % of GDP). It included promises to address global warming and to reverse the trend of growing inequality. The $3.6 trillion Federal budget proposal passed the House on April 2, 2009 by 233–196, with even “blue dog” conservative Democrats supporting it, but, again, no Republicans.

Obama’s social policies even received a modicum of success; on January 22, 2009 a bill against pay discrimination passed the Senate 61–36. The House also gave final approval on February 4, by a vote of 290–135, to a bill extending health insurance to millions of low-income children. Forty Republicans voted for the bill, and 2 Democrats voted against it. When the bill was signed by President Obama, it was seen as the first of many steps to guarantee health coverage for all Americans but it was not clear that the battle over broader healthcare legislation would take most of 2009.

Obama gained another important victory when the Senate confirmed Sonia Sotomayor as Supreme Court Justice on August 6, 2009, by a vote of 68–31. She is the first Hispanic and the third woman to serve on the Court. Similarly, Obama nominated another woman, Elena Kagan, to the high court and she was confirmed almost exactly one year after Sotomayor on August 7, 2010 by a vote of 63–37. Though adding two left-leaning female justices to the court has increased the number of women on the Supreme Court to an all time high of 3, it has not fundamentally changed the ideological make-up of the current court which still regularly splits 5–4 in favor of more right-leaning rulings.

Events in 2009 and 2010 are consistent with the model presented in Schofield and Miller (2007). Obama is attempting to attract and retain pro-business social liberals with his response to the economic crisis while his massive budget proposal addresses the economic down-turn but has angered most Republicans. It is possible that the Republican Party will gain votes from the blue-collar voters who are suffering the most from the economic collapse. However, if there is any economic recovery by the 2012 election, it is possible that many of the pro-business groups in the country will respond to Obama’s attempt to get the economy moving by supporting him. That could leave the Republican Party with nothing but the old-style populism of William Jennings Bryan: anti-Wall Street, anti-banking, anti-Detroit, anti-immigration, and pro-evangelical religion. This will result in a party realignment to a situation where the socially liberal and economically conservative “cosmopolitan” Democrats are opposed to populist Republicans.[22]

In October, 2009, one group identifying as populist Republicans, the “Tea Party” activists opposed Obama’s policies on health care so much that they began lining up against the centrist Governor Charlie Crist in the GOP Senate primary. Ultimately, Crist was forced to become an Independent and a Tea Party darling, Marco Rubio, was nominated as the GOP candidate for the Florida Senate seat (and ultimately won the seat, beating Crist handily). Similarly, on November 1, 2009 the centrist Republican candidate, Dede Scozzafava, decided to drop out of the special election in New York’s 23rd congressional district and endorse the

[22] That is, unlike the situation in the previous figures, the Republican Party will move to the lower left quadrant of the policy space, while business interests in the upper right quadrant will switch to the Democrats. It is indicative of this trend that on April 28 Arlen Specter, the senator from Pennsylvania, shifted his allegiance from the Republican Party to the Democrats.

Democrat candidate, Bill Owens. Owens won the election in a district that had been Republican since 1872.

As the Healthcare debate heated up over summer and fall of 2009 it became clear that Republicans were intending to continue utilize their blocking coalition as long as possible to stimy Obama and the Democrats. Interestingly, some Democrats contributed to this opposition as well; in the health bill vote in the House in early November 2009, 219 Democrats with 1 Republican voted for the bill, while 176 Republicans and 39 "Blue Dog" Democrats voted against.[23] By December 19, Senator Bernie Sanders of Vermont, an independent who caucuses with the Democrats, as well as Democrat Senators Ben Nelson and Sherrod Brown, had agreed to a compromise bill. This brought the size of the coalition to the critical size of 60 votes, sufficient to force a decision in the Senate.[24] Finally on Christmas Eve, 2009, the health bill passed in the Senate, again by 60 votes with 39 Republicans opposed. However, the victory by Republican Scott Brown in the special Senate election in Massachusetts on January 19 deprived the Democrats of the 60 seat majority required to push through the legislation. On February 25, 2010, an attempt to reach a bipartisan compromise failed, and there was talk of using a manoeuvre known as "reconciliation" to force though a health bill using simple majority rule.[25] Finally, on March 25, after strenuous efforts by President Obama and House speaker, Nancy Pelosi, the House voted 220–207 to send a health care bill to the President. Republicans voted unanimously against the legislation, joined by 33 dissident Democrats. The Senate passed the bill by simple majority of 56–43, as required under reconciliation and the President signed a draft of the bill, the "Patient Protection and Affordable Care Act," on March 23, 2010 and an updated version of the bill on March 30, 2010.

While it seemed that "gridlock" ensued over the health care legislation, several other major pieces of legislation passed with far less opposition. On February 22, 2010 and again on March 17, 2010 the Senate voted 62–30 and 68–29 respectively to implement two multi-billion "jobs creation" programs. Even though the vote to end debate on the Financial Regulation bill failed to obtain the required supramajority on May 19, 2010, it eventually passed the Senate. On July 15, 2010 the Senate voted 60–39 for the Dodd-Frank Wall Street Reform and Consumer Protection Act, and this was signed into law by President Obama on July 21.[26]

[23] On Saturday, November 21, the Senate voted 60–40, along partisan lines, to move to the final discussion on the health care bill.

[24] Cloture is a motion aimed at bringing debate to an end. It originally required a two-thirds majority, but since 1975 has required a super-majority of 60.

[25] Reconciliation is a measure whereby a bill can pass the Senate with a simple majority; the legislation must be shown to be budget neutral over a ten-year span in accordance with the Byrd rule.

[26] This complex bill was 2300 pages long. Russ Feingold, a Democrat, voted against the bill, because it was not strong enough. Three moderate New England Republicans, Snowe and Collins of Maine, and Scott Brown of Massachusetts, voted for the bill. The death of Senator Robert Byrd of West Virginia made it more difficult to summon the required 60 votes for cloture.

President Obama also signed into law a bill to restore unemployment benefits for millions of Americans who have been out of work for six months or more.

Further complicating issues of partisan discontent in Congress has been the introduction of ever increasing quantities of money in the American political system. For example, in 2009, health care, pharmaceutical and insurance lobbyists[27] spent approximately $650 million on lobbying itself, and about $210 million on media advertising, while the oil and gas industry spent about $560 million.[28] It would seem inevitable that the importance of lobbying can only increase in the future.[29] The Supreme Court decision, *Citizens United versus Federal Election Commission*, on January 21, 2010, removed limits on campaign contributions and will further increase the importance of activist contributions. An earlier Court decision, *Federal Election Commission versus Wisconsin Right to Life Inc.* had allowed corporations to buy advertisements supporting candidates as long as they did not appeal explicitly for the election or defeat of a particular candidate. *Citizens United* removed this limitation.

In his State of the Union address in late January, 2010 President Obama said the court had "reversed a century of law that I believe will open the floodgates for special interests—including foreign corporations—to spend without limit in our elections."[30] Dworkin (2006) later called the Supreme Court decision "an unprincipled political act with terrible consequences for the nation."

In July, 2010, the Federal Election Commission approved the creation of two "independent" campaign committees, one each from the left and right, expressly designed to take advantage of the lack of spending limits. One committee is being set up by the Club for Growth, the conservative advocate for low taxes and less government. The other, called Commonsense Ten, with close ties to the Democrats, will raise money from individuals, corporations and unions. Both groups will be able to spend unlimited amounts, thanks to the *Citizens United* decision. A Democrat effort to impose new campaign finance regulations before the November congressional election was defeated on July 27 when all 41 Senate Republicans blocked a vote on a bill that would force special interest groups to disclose their donors when purchasing political advertisements. A second attempt at cloture on the bill failed by 59–39 in the Senate on September 23.[31]

As the 111th Congress drew to a close in November, 2010 there remained four major pieces of legislation on the agenda: A Deficit Reduction Act, an Expanded

[27] The pharmaceutical industry was a strong supporter of reform of health care, because of an agreement with Obama to protect the industry's profits.

[28] Tomasky (2010) gives a figure of $3.47 billion for spending by lobbyists in the non election year of 2009, citing data from the Center for Responsive Politics.

[29] Indeed, Herrera et al. (2008) observe that spending by parties in federal campaigns went from 58 million dollars in 1976 to over 1 billion in 2004 in nominal terms.

[30] Notably, George W. Bush appointed Supreme Court Justice Samuel Alito broke from traditional judicial decorum at State of the Unon speeches to shake his head in disagreement with the President reportedly muttering the words "that's not true."

[31] As usual it required 60 votes.

Trade and Export Act, a Comprehensive Immigration Act, and an Energy Independence and Climate Change Act. Despite passage by the House on June 26, 2009, the Waxman-Markey climate change bill, formally called the American Clean Energy and Security Act (ACES), never reached action in the Senate. On July 22, 2010, the effort to push forward with the Climate Change Act collapsed due to Republican opposition to a carbon tax. If these bills continue to prove impossible to enact because of partisan strife and opposition, the electorate is likely to oppose any incumbent due to their lack of efficacy at passing key legislation.

Given these uncertainties surrounding policy choices in the legislature, it is hardly surprising that voters in the United States doubt that government can be effective. Part of the problem would appear to be the degree of political polarization that results from the power of interest groups located in the opposed quadrants of the policy space.

5.1 Implications of the 2010 Election

In the November, 2010 mid-term election large amounts of money were funnelled through non-profit advocacy groups that can accept unlimited donations and are not required to disclose their donors. As of November 1, 2010, it was estimated that these groups had spent $280 million, 60 % from undisclosed donors. Three activist groups, the US Chamber of Commerce, American Crossroads and the American Action Committee had spent $32.8, $21.6 and $17 million respectively.

Former Bush advisers, Karl Rove and Ed Gillespie, first formed American Crossroads as a 527 independent-expenditure-only committee, but was required to disclose donors. They then formed Crossroads Grassroots Policy Strategies (GPS) as a 501(c)(4) social welfare nonprofit, which means it does not need to disclose donors, but is not supposed to be used for political purposes. GPS spent $17 million. The Chamber of Commerce is a 501(c)(6) nonprofit, but corporations that donate to the Chamber must disclose these contributions in their tax filings. These corporations include Dow Chemical, Goldman Sachs, Prudential Financial and the most highly publicized was a singular donation in excess of $1 million from Rupert Murdoch's News Corporation.

In addition to the external activist groups, South Carolina Senator, Jim DeMint, used the Senate Conservatives Fund as a PAC to funnel about a $1 million to many of the most right-wing Tea Party candidates. Indeed, a key element of the successful Republican campaign was that these activist bodies were able to target House and Senate races where incumbent Democrats were weak.

In the 2010 election cycle total campaign spending was about $4 billion, with Republican spending somewhat higher than total Democrat spending. The extremely high level of expenditure (especially for a midterm election) is particularly of interest because there is evidence that the policy positions of activists on the social axis has become more polarized over the last forty years (Layman et al. 2010). This polarization appears to have benefited the wealthy in society and may well account for the increase the inequality in income and wealth distribution that has occurred (Hacker and Pierson 2006, 2010; Pierson and Skocpol 2007).

Ultimately, the electorate seems to have blamed incumbents, particularly Democrats, for economic woes. In the midterm election in November, 2010, the Democrats lost 63 seats in the House, leading to a Republican majority of 242–192. In the Senate the Democrats lost 6 seats but retained a majority of 51–46 (with 3 Independents).[32] Many of the newly elected members of Congress received the backing of the Tea Party and vocally subscribed to extreme policy stances like abolishing the Federal Reserve, unemployment benefits, and even income taxes. Further, preliminary demographic studies of the Tea Party indicate that they are predominantly older, middle class suburban and rural white Americans.[33] This demographic make-up leads one to postulate that the Tea Party is a representation of a populist movement supported primarily by elites in the South and West. Although tea party supporters are opposed to deficit spending, they generally are supportive of social security and medicare, and want to reduce the deficit by cutting other programs. Perhaps most striking about the Tea Party is the immediate impact they had on Congress itself; the Republican House leadership even created a special leadership post for a Representative from the Tea Party wing.

Because of the plurality nature of the U.S. electoral system, parties have to build a winning coalition of mobilized disaffected activists and current party activists (Miller and Schofield 2003). Many of the tea party activists see themselves as conservative independents that are opposed to big business despite the fact that large corporations and wealthy individuals heavily funded many of the tea party candidates campaigns. Even before the 112th Congress entered session the Republican Party stood up for the wealthy benefactors by insisting on blocking all legislation during the lame duck session until the wealthiest two percent of Americans received the same extension on their tax cuts that the other 98 % were set to receive. This Republican measure included blocking discussion on repealing the "Don't Ask, Don't Tell" legislation, immigration reform legislation, a nuclear arms treaty and even legislation allocating funds to provide healthcare to September 11, 2001 first responders.

In an effort to close his career with parting advice about compromise retiring Connecticut Senator Chris Dodd gave his valedictory speech on the Senate floor on November 30, 2010 with remarks including the following:

> From the moment of our founding, America has been engaged in an eternal and often pitched partisan debate. That's no weakness. In fact, it is at the core of our strength as a democracy, and success as a nation. Political bipartisanship is a goal, not a process. You

---

[32] This was the backlash predicted by Bunch (2010). However, the Democrat losses may be due to the spending pattern. The *New York Times* analysis suggested that in 21 House districts where groups supporting Republican candidates spent about $2 million, they won 12.

[33] Skocpol and Williamson (2010) have been collecting survey and interview data on the Tea Party since its emergence and all indications are that Tea Party members are a very specific demographic sub-group with traditional populist concerns. See also Rasmussen and Schoen (2010).

> don't begin the debate with bipartisanship—you arrive there. And you can do so only when determined partisans create consensus—and thus bipartisanship. In the end, the difference between a partisan brawl and a passionate, but ultimately productive, debate rests on the personal relationships between Senators.

Another elder statesman in the Senate, Indiana's Richard Lugar, clearly felt the same way as Senator Dodd after the 2010 election as he defied the Republican Party over their various demands. Senator Lugar has said that the environment in Washington was the most polarized he had seen since joining the Senate in 1977. John C. Danforth, the former Republican senator from Missouri, remarked that

> If Dick Lugar, having served five terms in the U.S. Senate and being the most respected person in the Senate and the leading authority on foreign policy, is seriously challenged by anybody in the Republican Party, we have gone so far overboard that we are beyond redemption.

President Obama eventually struck a deal to allow the tax cuts to be extended for all Americans (in exchange for an extension of unemployment benefits) despite the fact that even the most positive economic forecasts do not predict these tax cuts to the wealthy bringing unemployment down by more than 0.1% over the two year lifespan of the tax cut extension.This compromise angered many in the liberal wings of Democratic Party as they saw compromise as a betrayal of President Obama's progressive values. In the wake of persistent attack by several prominent liberal Democrats, Obama invited former President Bill Clinton to give a White House press conference in support of the compromise. Involving the former President in this way can be seen as either an act of desperation or an attempt by the administration to harken back to the 1990s (or earlier) when compromise was an acceptable political tactic.[34]

On Monday December 13, 2010 the Republican bargaining ploy worked. The Senate voted to halt debate on the tax cut bill. Other provisions of the $858 billion bill would extend unemployment insurance benefits and grant tax breaks for schoolteachers, mass transit commuting expenses and landowners who invest in conservation techniques. The compromise bill overwhelmingly passed the Senate on December 15 by a vote of 81–19. Despite accusations by House Speaker, Nancy Pelosi that Republicans were forcing Democrats "to pay a king's ransom in order to help the middle class" at midnight on December 16 the measure passed with 139 Democrats and 138 Republicans in favor and 112 Democrats and 36 Republicans opposed. President Obama signed the bill into law the next day.

After this initial compromise was struck, the logjam seemed to have broken as Congress began debate on repealing "Don't Ask, Don't Tell," on the passage of the nuclear arms treaty, and on temporary measures to continue funding the federal government into 2011. This step toward compromise and productivity irked Senators Jon Kyl (Republican from Arizona) and Jim DeMint (Republican from South Carolina) who criticized Majority Leader Harry Reid (Democrat from

[34] It is worth noting that the Founding Fathers repeatedly cited the need for compromise as one of the greatest strengths of the U.S. political system.

Nevada) for "disrespecting" the institution and the Christmas holiday by putting so much work on the Congressional docket that Senators might need to return to work during the week between Christmas and New Year. These statements by Senators Kyl and DeMint provide a stark reminder of the roadblocks to compromise in activist driven politics. House and Senate Republicans derailed a $1.2 trillion spending measure put forward by Senate Democrats, and promised to use their majority in the new House to shrink government. On December 21 Congress did approve a temporary spending bill up until March 2011.

On December 18, the "Dream Act" to allow illegal immigrant students to become citizens failed on a Senate vote of 55–41, but the Senate did vote 65–31 to repeal the "Don't Ask, Don't Tell" legislation, making it possible for gays to serve openly in the military. The House had previously approved this repeal by 250–175.

On December 20, the Senate voted 59–37 to reject an amendment to the new arms control treaty, New Start, with Russia. The amendment would have killed the treaty because any change to the text would have required the United States and Russia to renegotiate the treaty. Two days later the Senate voted 71–26 for the treaty. This treaty was seen as the most tangible foreign policy achievement of President Obama. Thirteen Republicans joined a unanimous Democratic caucus to vote in favor, exceeding the two-thirds majority required by the Constitution.

As Obama said:

> I think it's fair to say that this has been the most productive post-election period we've had in decades, and it comes on the heels of the most productive two years that we've had in generations. If there's any lesson to draw from these past few weeks, it's that we are not doomed to endless gridlock. We've shown in the wake of the November elections that we have the capacity not only to make progress, but to make progress together.

Given the results of the 2010 elections and the fact that increasingly the Democrats in Congress represent the richest and the poorest constituencies, while the Republican Party is no longer the party of the wealthy but that of the disillusioned middle class and the ultra-wealthy, the indications for the 112th Congress are that, with a divided Congress and increasingly activist driven politics, conflict between the two parties will not only continue but escalate in the run up to the 2012 election.

One of the first moves by the House in the new 112th Congress was to vote, on January 19, 2011, to repeal the Health Care Bill by a margin of 245–189. However, this repeal cannot pass the Democratic majority in the Senate.

A shutdown of government in early April, 2011, was only just averted by a compromise that cut the budget by $38 billion. After much wrangling, the House passed legislation on April 14, to finance the federal government for the rest of the fiscal year. The final House vote was 260–167, with 59 members of the House Republican majority and more than half the Democratic minority voting against the legislation. The bill also passed the Senate 81–19, again with many Republicans opposed. On April 15, the House voted 235–193 to approve the fiscal blueprint for 2012, drafted by Representative Paul D. Ryan, Republican of Wisconsin and chairman of the Budget Committee. The blueprint proposed a cut in expenditure of $5.8 trillion over the next decade.

By July, it seemed that the political system was again in gridlock with the parties completely polarized over the question of the US public debt. The debt ceiling was at $14.3 trillion and the current US Treasury debt was $14.29 trillion.[35] Republicans demanded a reduction in spending and the maintenance of tax cuts, while Democrats basically wanted the opposite, continued spending on social programs and tax increases on certain segments of the population.

The House on Friday July 29, finally approved a plan for a short-term increase in the debt ceiling and cuts in spending. The vote was 218–210, with 22 Republicans unwilling to support the efforts by House Speaker, John A. Boehner, to get a bill approved. This ended a week of intense fighting among Republicans. The game then shifted to the Senate which tabled the House proposal. On August 1 the House of Representatives passed a compromise bill, 269–161, supported by Democrats, increasing the debt ceiling by $400 billion, with an additional $500 billion through February, with spending caps of over $900 billion. A newly designed joint committee was vested with the responsibility of determining future cuts of over $1 trillion. The Senate passed the bill 74–26 on August 2 with 19 Republicans, and 6 Democrats and one independent voting against. President Obama immediately signed the bill into law. Despite the eventual compromise on the debt ceiling, on August 5, 2011 Standard and Poor, the credit rating agency, downgraded US Federal debt from AAA to AA+, and the Dow industrial index dropped about 20 % in the following days. However, demand for U.S. Treasury Bonds increased.

Later in August the 2012 Republican Presidential primary season began. Early frontrunners included Tea Party darlings Representative Michele Bachman, Representative Ron Paul, and Governor Rick Perry. Former Governor Mitt Romney openly admitted seeking a more centrist route to the nomination but he will have to contend with activist money such as the PAC "Make Us Great Again" which plans on supporting Rick Perry to the tune of $55 million.

On September 13, President Obama acted on the economic turmoil set off by the Debt Ceiling debate, Standard and Poor downgrade and continuing European debt crisis by sending a $447 billion jobs bill to Congress. Initial reaction from Republicans indicated a willingness to accept some measures of the bill, coupled with an insistence on keeping tax cuts for the wealthiest and resistance to closing corporate loopholes. On November 21, however, the Committee to reduce the deficit announced that it could not come to any agreement, followed by the remark "We remain hopeful that Congress can build on this committee's work and can find a way to tackle this issue in a way that works for the American people and our economy." The Dow closed about 2 % down for the day.

---

[35] Of this $6.2 trillion is held by the US government, $2.7 trillion in the Social Security Trust Fund, $1.9 trillion in other government agencies and $1.6 trillion in the Federal Reserve. China and Hong Kong hold $1.3 trillion, other countries hold $3.2 trillion, the remaining $3.6 trillion is held by pension funds etc.

The debate over the jobs bill highlights the fact that, despite media attention to the contrary, Obama has attempted to attract and retain pro-business social liberals with his response to the economic crisis. In addition to naming General Electric CEO Jeffrey Immelt as Chairman of the President's Council on Jobs and Competitiveness, the President's second Chief of staff is former Commerce secretary and bank executive William Daley. These steps, along with his massive budget proposals providing relief to banks and other businesses in order to address the economic down-turn, has angered many in populist circles. Meanwhile, insistence on closing corporate tax loopholes and the spectre of increased financial regulation, has eroded business support for the President.

This lack of support in both the populist and cosmopolitan quadrants leaves the President and his party vulnerable to attacks by traditionally conservative Republicans as well as to the more populist demands of the Tea Party. As a result of persistently high unemployment rates, populist anger has spiked and even spawned a second, distinctly liberal-minded populist group, the "Occupy Wall Street" protesters. Given how amorphous this groups interests are, as of this writing they have been unable to garner much support from mainstream U.S. politicians but they have begun to receive a great deal of media attention causing several dozen protests to spring up around the U.S. as well as Europe. So, barring a great increase in political clout by the "occupy Wall Street" crowd it is possible that the Republican Party will continue to gain votes from the blue-collar voters who are suffering the most from the economic collapse. Should the Republican party cater to the traditional populist demands expressed by those in the Tea Party, they will be hearkening back to an era of old-style populism as expressed by William Jennings Bryan: anti-Wall Street, anti-banking, anti-Detroit, anti-immigration, and pro-evangelical religion. This will result in a party realignment to a situation where the socially liberal and economically conservative "cosmopolitan" Democrats are opposed to populist Republicans. That is, the Republican Party may begin to move to the lower left quadrant of the policy space, while some business interests in the upper right quadrant will switch to the Democrats.[36] Unlike the situation in Fig. 1, over the long term, the partisan cleavage line may rotate further in a clockwise direction.

6 Land, Capital and Labor in U.S. History

The activist model presented in this chapter can be used to explain the conflict of land and capital that dominated US politics in the nineteenth and early twentieth centuries and to some extent, still persists today.

Schofield (2006b) argues that Britain's ability to fight the long eighteenth century war with France depended on a compact between land and capital that was

[36] For example, on April 28, 2010 Arlen Specter, the Senator from Pennsylvania, shifted his allegiance from the Republican Party to the Democrats.

put in place by Robert Walpole, in the 1720s, and lasted until the repeal of the Corn Laws in 1846. The compact was based on the protection of the agrarian interest by customs and excise, and required the disenfranchisement of most of the population until the First and Second Reform Acts of 1832 and 1867.[37]

The Declaration of Independence by the thirteen colonies in 1776 was, in turn, triggered by conflict over land, specifically because of the attempt by the British to remove the Ohio Valley from settlement though the Quebec Act of July 1774. This Act led almost immediately to the First Continental Congress in October 1774, and was denounced in the Declaration itself.

In the United States after independence, conflict between Federalists, represented particularly by Alexander Hamilton, and the Republicans, James Madison and Thomas Jefferson, focused on capital versus land. Hamilton's Reports of 1790–1791 on Public Credit, Manufactures and The National Bank were all aimed at creating an American analogue of the British system of tariffs and excise. Since the United States exported land-intensive goods, the only feasible path to creating a commercial economy was to sustain manufactures either by tariff or by direct government assistance. Hamilton rejected the Madison-Jefferson view that the future of the U.S. economy lay principally in the cultivation of the land. Indeed, in the Report on Manufactures, Hamilton argued that the U.S. could grow only through an increase of productivity as a result of manufacturing.

Madison and Jefferson believed that Hamilton's commercial empire in the United States would generate precisely the same phenomenon of immoderation and disenfranchisement as had occurred in Britain. Hamilton's scheme would mean tariffs to raise revenue, increasing government debt, an extensive military establishment and corrupt "placemen." Jefferson's "Empire of Liberty" meant the exact opposite[38] and his election in 1800 saw the victory of the Democrat-Republican trade-oriented coalition of the slave-owning elite and free agrarian labor against the more urban north east.[39]

Until the election of Lincoln in 1860. the political coalition structure was "intersectional" between the eastern Whigs and the agrarian Democrats of the south and west. Lincoln's election was the result of the collapse of the agrarian coalition largely triggered by the Dred Scott opinion of the Supreme Court in 1857. Lincoln argued that this decision could lead to the expansion of slavery to the Pacific, against the interests of northern free labor.

---

[37] The 1867 Act was the most extensive. See McLean (2001); Acemoglu and Robinson (2000) for discussion.

See also Acemoglu and Robinson (2006a, b) for a discussion why Great Britain's path to economic development was not blocked by agrarian elites in this period.

[38] See the discussion of this period in Wood (2009)

[39] In this election, the Democrat-Republicans won 146 electoral college votes, with Jefferson and Burr, of New York, each receiving 73. The Federalists won 129 in total. Eventually Jefferson won the House with ten states to four for Burr. The three fifths weight given to unfree labor in the south had proved crucial.

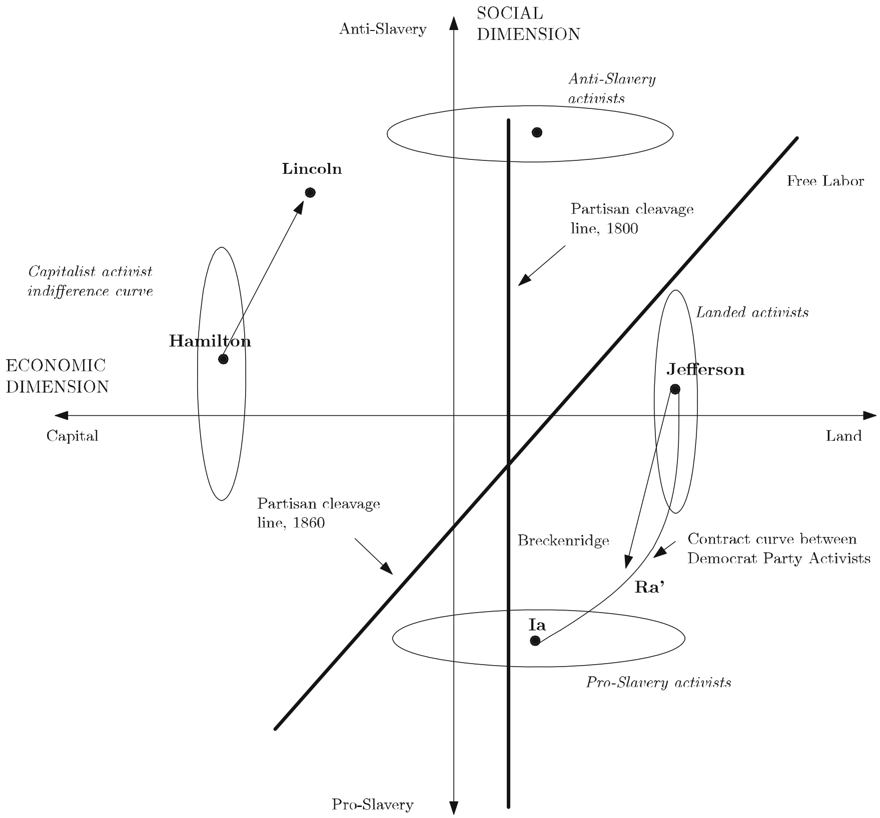


Fig. 8 Changes in political realignment 1800–1860

Figure 8 gives a heuristic representation of the transformation in party positions between the election of Jefferson in 1800 and the onset of the Civil War.

During the Civil War, the Tariff Acts of 1862 and 1864 were proposed as means to raise capital for the effort against the south, but as Taussig (1888) noted, in his classic treatise on the tariff,

> Great fortunes were made by changes in legislation urged and brought by those who were benefited by them.

By the Tariff Act of 1883, the average duty on aggregate imports was of the order of 30 %, mostly on manufactures.

The second half of the nineteenth century had seen an enormous growth of agrarian exports from the U.S to Great Britain. As Belich (2009) notes, grain exports increased from a million tons in 1873 to 4 million by 1900, with similar increases in dairy and meat products. However, by 1900, the “Dominions” (Canada, New Zealand and Australia) began to replace the United States as the agrarian suppliers for Britain. At the same time, the United States began its somewhat delayed process of industrial development, making use of the transport infrastructure, canals etc that had been put in place in the previous decades. Belich (2009) suggests that the decoupling of the United States from Britain took place about 1900, by which time the population of New York had reached 3.5 million.

This decoupling sets the scene for the conflict between the manufacturing interests of the north-east, and their preference for the protective tariff, against the free trade preference of the south and west of the country at the election of 1896. In this election Republican William McKinley stood for the manufacturing interests and barely defeated the Democrat, William Jennings Bryan whose populist position for cheap money against the gold standard was strongly supported in the somewhat less populous agrarian south and west.[40] Figure 9 again gives a representation of the realignment between 1860 and 1896, while Fig. 10 continues with the realignment as Wilson shifted to a position in the upper left quadrant of the political space. F.D. Roosevelt in the 1930s continued with this realignment.[41]

The Smoot-Hawley Tariff Act of 1930 raised average tariffs to about 20 % and is generally considered to have contributed to the dramatic fall in both imports and exports. By 1993, however, the massive economic growth of the post war years led to the North American Free Trade Agreement, in 1993, pushed forward by William Clinton. Even though populists, like Patrick Buchanan (1998) have hated the resulting globalization, it contributed to the period of rapid growth that came to such an abrupt end recently.[42]

This continuing realignment has changed the *heartland* of each of the two parties. In the late nineteenth century, the north-east was industrial and strongly Republican. The rest of the country was agrarian and Democrat. By the early part of our new century, the north-east was socially liberal and Democrat, while the rest of the country was basically socially conservative and Republican.

In recent years much discussion has focused on why North America was able to follow Britain in a path of economic development, but Latin America and the Caribbean islands, though generally far richer initially, fell behind in the nineteenth century. In their discussion of Latin American economic development, Sokoloff and Engerman (2000) have emphasized the different factor endowments of North and South America. In contrast, Przeworski and Curvale (2006) argue that while economic inequality tended to persist and has been related to the degree of political inequality, many aspects of the developmental path appear highly contingent. Indeed whether Latin American economies grew, and the extent to which they have protected the factors of capital and labor, seems to be dependent on shifting balances of power between differing activist groups, as suggested by the formal model presented in this chapter.

---

[40] McKinley won 51 % of the popular vote but 60 % of the electoral college, taking the entire northeast along with California and Oregon.

[41] See also Schofield et al. (2003).

[42] As at the end of the nineteenth century, the recent period in the U.S. has been characterized by increasing income inequality. According to the Economic Policy Unit, the top 1 % of Americans currently own 34 % of the net worth of the country.

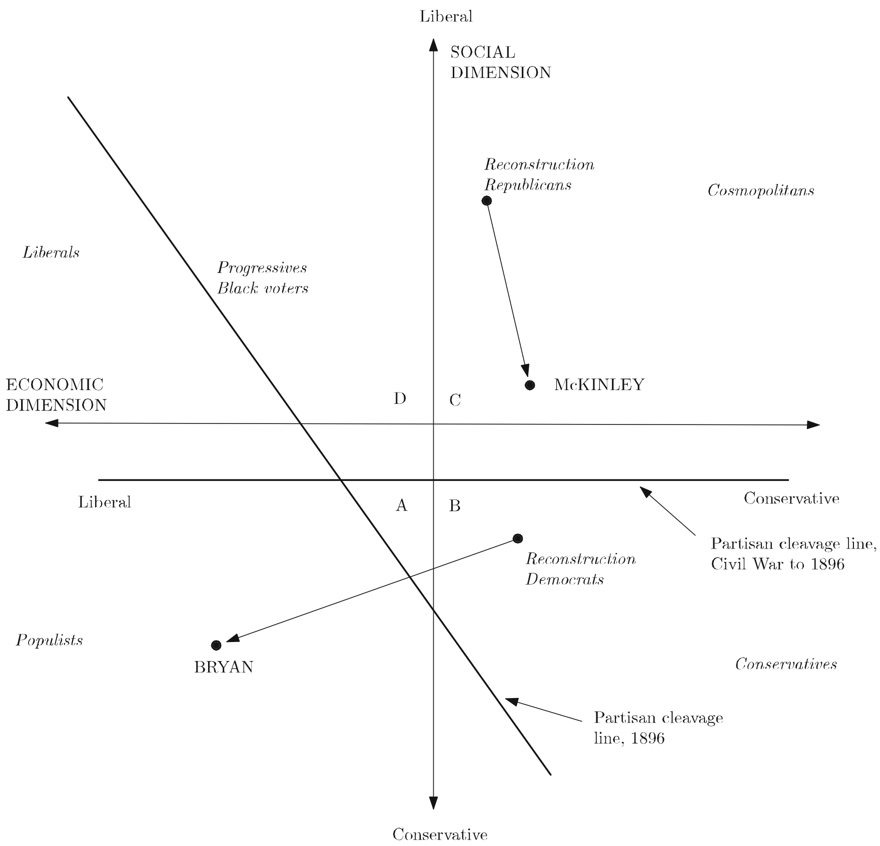


Fig. 9 US realignments 1860–1896

Galiani et al. (2012) have applied a variant of the model presented here to elucidate the conflicts that exist between activist groups which are characterized by their control of different economic factors. They argue that Latin American economies are *diversified natural resource-rich economies*, which tend to have an important domestic industry that competes with the imports. In such a political economy parties tend to diverge and trade policy is likely to be more protectionist and unstable. They suggest that uncertainty in policy has been one cause of the slower development path of these economies. In principle this extended model can incorporate activist and citizen preferences over levels of trade protection[43] and moves towards democratization.

Acemoglu et al. (2008, 2009) discuss the hypothesis of "critical junctures" in discussing moves to democracy. Such a notion parallels that of uncertainty over

[43] See Rogowski (1989).

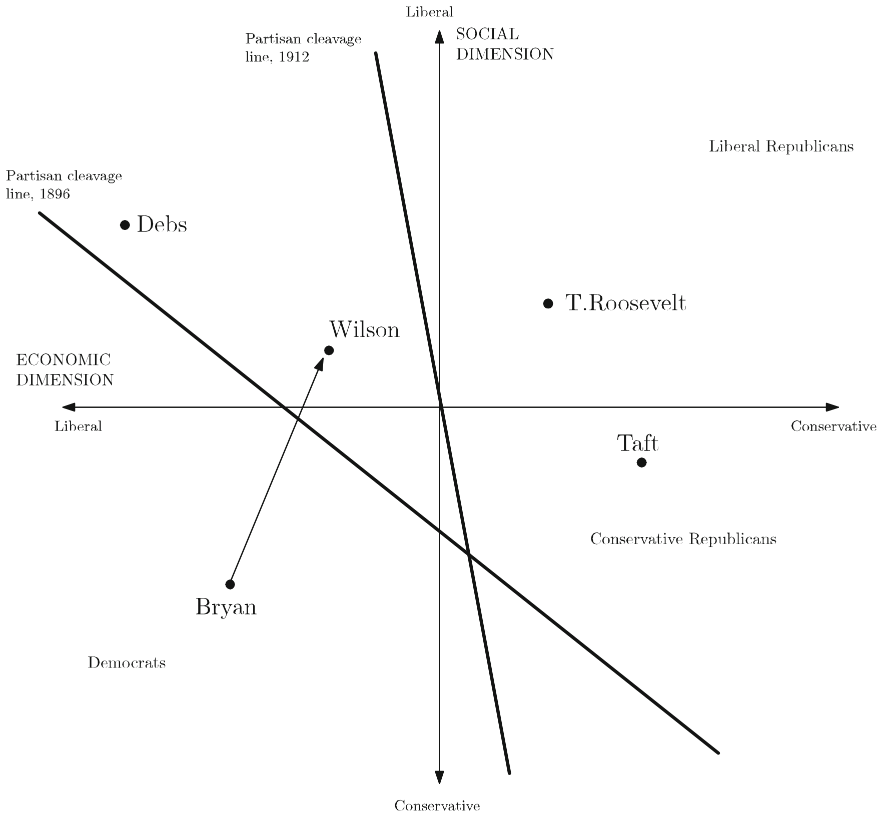


Fig. 10 The election of 1912

the nature of the various elite activist coalition that must choose whether to support the autocrat or reformers.[44]

This brief sketch of shifts in the dominant societal cleavages indicates how social choice in both developed and less developed polities will tend to be transformed as a result of essentially political changes in the balance of power between landed and capital elites in coalition with different elements of enfranchised labor. As the Tea Party has shown in 2010, various elites, primarily in the south and west, have successfully mobilized against entrenched capital elites (largely cosmopolitans) in the wake of the economic crisis. These "landed" elites have mobilized socially conservative labor, especially older middle class labor, to vote for the GOP. Disillusioned young labor and discouraged capital elites failed to turn out for Democrats, leading to large Republican gains.[45] Due to the

[44] Schofield and Levinson (2008) have applied an early version of the model here to discuss the collapse of autocracies in Argentina, Franco's Spain and the Soviet Union.

[45] Youth voter turnout declined substantially from 2008 to 2010.

economic crisis and President Obama's frequent populist tone, capital elites (i.e. bankers) did not fund Democrats in 2010 to the levels they did in 2008.

The ultimate compromise between land and capital in the U.S. occurred in 1787 when the Senate was created to appease states with small populations, like Delaware. This "Great Compromise" still deeply influences United States politics. In the 111th Congress the Democrats in the Senate represented more than 63% of the United States population but held only 57 of the 100 Senate seats.[46] The result was months spent stuck in gridlock over healthcare legislation and many other pieces of the legislative agenda left to die in committee. The Democrat's super-majority was insufficient to overcome the filibuster precisely because each state receives two senators, regardless of the state's population.[47] In a sense, the landed elite in the U.S. has currently won a skirmish with the capital elite because of a constitutional decision made more than two hundred years ago.[48]

7 Concluding Remarks

The volatility of recent elections in the United States has provided a window into how democratic elections can lead to extremely non-convergence behavior. Activist valence has also played an increasingly large role in U.S. elections of late, especially since the *Citizen's United* Supreme Court decision in January of 2010. It is increasingly apparent that the increased polarization that has led to turnover in Congress. Volatility in American politics is a natural result of a system developed more than two hundred years ago on a basic premise that political parties would not play a role in American politics. Given this background, this chapter has applied a theoretical stochastic model to present a discussion of recent elections in the United States. We have also applied the model to earlier realignments in the fundamental political configuration as the economy shifted to manufacturing in the late nineteenth century and early twentieth century. The model also allows us to contrast the situation in the 1960s with the present.

After Kennedy was elected President in 1960 (by a very narrow margin of victory against Nixon), he delayed sending a Civil Rights Bill to Congress,

[46] This is calculated by determining the population of the state and the party of the Senators that represent that state. In the 112th Congress the Democrats represent more than 60 % of the U.S. population but hold only 53 of the 100 seats in the Senate. These calculations do not include Washington D.C. which does not have representation in the Senate.

[47] As of 2010 Wyoming had a smaller population than Washington D.C. but Wyoming continues to have two Senators and a Representative while Washington D.C. has only a non-voting Representative.

[48] Acemoglu and Robinson (2006b) discuss the attempts by agrarian elites in countries such a Russia and Austria- Hungary in the nineteenth century to resist industrialization. There may be an element of similar resistance by certain elites in the U.S. to the transformation to an advanced idea dependent economy of the kind discussed in Jones (2002, 2009)

precisely because of the possible effect on the South (Branch 1998). To push the Civil Rights Act through in 1964, Johnson effectively created, with Hubert Humphrey's support, an unstable coalition of liberal northern Democrats and moderate Republicans, with sufficient votes in the Senate to effect 'cloture', to block the southern Democratic filibusters.[49] This was the first time since Reconstruction that the Southern veto was overwhelmed. The danger for Johnson in the election of 1964 was that a Republican candidate could make use of the fact of Republican party support for civil rights to attract disaffected social liberals. Traditional Republican Party activists were thus in an electoral dilemma, but resolved it by choosing the southern social conservative, Goldwater.

The present gridlock between the legislative and executive branches is more extreme than in 1964 because there are now no moderate Republicans to join the social-liberal coalition. The electoral pivot line has rotated so that all Republicans are located in the socially conservative half of the policy space. In addition money has become more important and has made US politics "irrational". With money playing an increasingly large role in recent elections, this irrationality and non-convergence to the electoral center is likely to persist. Moreover, powerful activist groups in the cosmopolitan and populist sectors have the potential to draw in politicians and shift the partisan cleavage line between parties. Were it not for the resources the activist groups provide it would be irrational for politicians to move toward these activist bases . Simply put, the resources of economic activists further influence politicians so they cluster in opposing quadrants of the policy space.

Krugman (2012) argues that increasing inequality since the deregulation of the Reagan administration has led to the current dominance of money in the US political system. Sandel (2012) asks if there are "certain moral and civic goods that markets do not honor and money cannoy buy." it seems obvious that moral objections to the dominance of money in politics far outweighs the argument for "free speech," used to justify *Citizen's United*. Indeed there is a second economic argument. The electoral mean is a natural and socially efficient outcome of the political process, which would come to pass in the absence of money.

Popper (2008) argued that plurality electoral systems, otherwise known as "first past the post" were to be preferred to proportional electoral systems because they gave voters a clear choice. As we have seen, the constitutional structure of the US polity, coupled with the influence of money has recently tended to gridlock. Although there is the appearance of choice for the voters, government has been unable to come to grips with the severe quandaries briefly mentioned in the introduction. The absence of effective choice by the US increases uncertainty in policy-making thus creating a difficult situation for business and international leaders attempting to make long-term investments and policy decisions. Indeed, Posner and Vermeule (2011)

[49] Caro (2012): 568) describes the drama of the cloture vote of Jun 10, 1964 after a filibuster of 57 days with 27 Republicans and 44 Democrats voting aye. The bill passed on June 19 by 73–27. The voting Rights Act of 1965 passed again after a long fight by Johnston against Congress.

argue that the United States needs to reconsider its constitutional separation of powers in the presence of such gridlock and uncertainty.

On the other hand, the recent European debt crisis has led to the fall of governments in the multiparty systems of Ireland (February, 2011) Finland (2011), Portugal (June 2011), Denmark (September, 2011), Slovakia (October 2011), Greece, Italy and Spain (November 2011). Thus fragmented or proportional, multiparty systems, coupled with a fragile fiscal system based on the euro also seem to create difficulties in dealing effectively with the fall-out from the recession of 2008–2009.

References

Acemoglu, D., & Robinson, J. (2000). Why did the west extend the franchise? Growth, inequality and democracy in historical perspective. *Quarterly Journal of Economics*, *115*, 1167–1199.

Acemoglu, D., & Robinson, J. (2006a). *Economic origins of dictatorship and democracy*. Cambridge: Cambridge University Press.

Acemoglu, D., & Robinson, J. (2006b). Economic backwardness in political perspective. *American Political Science Review*, *100*, 115–131.

Acemoglu, D., & Robinson, J. (2008). Persistence of power, elites, and institutions. *American Economic Review*, *98*, 267–293.

Acemoglu, D., Johnson, S., Robinson, J., & Yared, P. (2008). Income and democracy. *American Economic Review*, *98*, 808–842.

Acemoglu, D., Johnson, S., Robinson, J., & Yared, P. (2009). Reevaluating the modernization hypothesis. *Journal of Monetary Economics*, *56*, 1043–1058.

Adams, J., & Merrill, S. (2005). Policy seeking parties in a parliamentary democracy with proportional representation: A valence-uncertainty model. *British Journal of Political Science*, *39*, 539–558.

Aldrich, J. H. (1983). A downsian spatial model with party activists. *American Political Science Review*, *77*, 974–990.

American national election studies 2000 and 2004.

Ansolabehere, S., Snyder, J. M., & Stewart, C. (2001). Candidate Positions in Congressional Elections. *American Journal of Political Science*, 45.

Aragones, E., & Palfrey, T. (2002). Mixed equilibrium in a downsian model with a favored candidate. *Journal of Economic Theory*, *103*, 131–161.

Ashworth, S., & Bueno de Mesquita, E. (2009). Elections with platform and valence competition. *Games and Economic Behavior*, *67*(1), 191–216.

Banks, J., & Duggan, J. (2005). Probabilistic voting in the spatial model of elections: The theory of office motivated candidates. In D. Austen-Smith & J. Duggan (Eds.), *Social choice and strategic decisions: Essays in honor of Jeffrey Banks*. Heidelberg: Springer.

Belich, J. (2009). *Replenishing the Earth: The settler revolution 1783–1939*. Oxford: Oxford University Press.

Branch, T. (1998). *Pillar of fire*. New York: Simon and Schuster.

Buchanan, P. (1998). *The great betrayal: How American sovereignty and social justice are being sacrificed to the gods of the global economy*. New York: Little Brown.

Bunch, W. (2010). *Backlash*. New York: Harper.

Calvert, R. L. (1985). Robustness of the multidimensional voting model: Candidates, motivations, uncertainty and convergence. *American Journal of Political Science*, *29*, 69–85.

Caro, R. A. (2012). *The passage of power*. New York: Knopf.

Coate, S. (2004). Political competition with campaign contributions and informative advertising. *Journal of the European Economic Association*, *2*, 772–804.
Condorcet, N. (1994). [1785] *Essai sur l'application de l'analyse a la probabilite des decisions rendus a la pluralite des voix*. Paris: Imprimerie Royale.
Coram, A. (2010). Resource spending over time in an electoral competition. *Electoral Studies*, *29*, 497–508.
Downs, A. (1957). *An economic theory of democracy*. New York: Harper and Row.
Duggan, J. (2006). Candidate objectives and electoral equilibrium. In B. R. Weingast & D. Wittman (Eds.), *The Oxford handbook of political economy*. Oxford: Oxford University Press.
Duggan, J., & Fey, M. (2005). Electoral competition with policy-motivated candidates. *Games and Economic Behavior*, *51*, 490–522.
Dworkin, R. (2006). *Is Democracy Possible Here?* Princeton: Princeton University Press.
Enelow, J. M., & Hinich, M. (1989). The location of American Presidential candidates. *Mathematical and Computer Modelling*, *12*, 417–435.
Erikson, R. S., & Romero, D. W. (1990). Candidate Equilibrium and the Behavioral Model of the Vote.*The American Political Science Review*, *84*(4), 1103–1126.
Fiorina, M., Abrams, S. J., & Pope, J. C. (2005). *Culture war?: The Myth of a polarized America*. New York: Pearson Longman.
Fiorina, M., & Abrams, S. J. (2009). *Disconnect: The breakdown of representation in American politics*. Norman: University of Oklahoma Press.
Galiani, S., Schofield, N., & Torrens, G. (2012). Factor endowments, democracy and trade policy divergence. *Journal of Public Economic Theory* (in press).
Groseclose, T., & Snyder, J. (1996). Buying supermajorities. *American Political Science Review*, *90*, 303–315.
Groseclose, T. (2001). A Model of Candidate Location When One Candidate Has a Valence Advantage.*American Journal of Political Science*. *45*(4), 862–886.
Grossman, G. M., & Helpman, E. (1994). Protection for sale. *American Economic Review*, *84*(1994), 833–850.
Grossman, G. M., & Helpman, E. (1996). Electoral competition and special interest politics. *The Review of Economic Studies*, *63*, 265–286.
Grossman, G. M., & Helpman, E. (2001). *Special interest groups*. Cambridge: MIT Press.
Hacker, J. S., & Pierson, P. (2006). *Off Center: The Republican revolution and the erosion of American democracy*. New Haven: Yale University Press.
Hacker, J. S., & Pierson, P. (2010). *Winner-take-all politics: How Washington made the rich richer-and turned its back on the middle class*. New York: Simon and Schuster.
Herrera, H., Levine, D., & Martinelli, C. (2008). Policy platforms, campaign spending and voter participation. *Journal of Public Economics*, *92*, 501–513.
Jackson, M., & Watts, A. (2010). Social games: matching and the play of finitely repeated games. *Games and Economic Behavior*, *70*, 170–191.
Jones, C. I. (2002). Sources of U.S. economic growth in a world of ideas. *American Economic Review*, *92*, 220–239.
Jones, C. I., & Romer, P. (2009). *The kaldor facts: Ideas, institutions, population and human capital*. Cambridge, MA: NBER working paper #15094.
Krugman, P. (2012). *End this depression now*. New York: Norton.
Layman, G. C., et al. (2010). Activists and conflict extension in American party politics. *American Political Science Review*, *104*, 324–346.
Madison, J. (1999, [1787]). The federalist No. 10. In J. Rakove (Ed.), *James Madison: writings*. New York: The Library of America.
McCarty, N., Poole, K., & Rosenthal, H. (2006). *Polarized America*. Cambridge: MIT Press.
McKelvey, R. D. (1979). General conditions for global intransitivities in formal voting models. *Econometrica*, *47*, 1085–1112.
McKelvey, R. D., & Patty, J. (2006). A theory of voting in large elections. *Games and Economic Behavior*, *57*, 155–180.
McLean, I. (2001). *Rational choice and British politics*. Oxford: Oxford University Press.

McLennan, A. (1998). Consequences of the condorcet Jury theorem for beneficial information aggregration by rational agents. *The American Political Science Review*, *92*, 413–418.

Meirowitz, A. (2008). Electoral contests, incumbency advantages and campaign finance. *Journal of Political Science*, *70*, 681–699.

Miller, G., & Schofield, N. (2003). Activists and partisan realignment in the U.S. *American Political Science Review*, *97*, 245–260.

Miller, G., & Schofield, N. (2008). The transformation of the republican and democratic coalitions in the United States. *Perspectives on Politics*, *6*, 433–450.

Penn, E. (2009). A model of far-sighted voting. *American Journal of Political Science*, *53*, 36–54.

Patty, J., Snyder, J. M., & Ting, M. (2009). Two's company, three's an equilibrium: Strategic voting and multicandidate elections. *Quarterly Journal of Political Science*, *4*, 251–278.

Peress, M. (2010). The spatial model with non-policy factors: A theory of policy motivated candidates. *Social Ch Welfare*, *34*, 265–294.

Pierson, P., & Skocpol, T. (2007). *The transformation of american politics: Activist government and the rise of conservatism*. Princeton: Princeton University Press.

Poole, K., & Rosenthal, H. (1984). U.S. presidential elections 1968–1980: A spatial analysis. *American Journal of Political Science*, *28*, 283–312.

Popper, K. (2008). The open society and its enemies revisited. *The Economist*, *307*, 19–22.

Posner, E. A., & Vermeule, A. (2011). *The executive unbound*. Oxford: Oxford University Press.

Przeworski, A., & Curvale, C. (2006). Does politics explain the economic gap between the United States and Latin America? In F. Fukuyama (Ed.), *La Brecha entre America Latina y los Estados Unidos*. Fondo de Cultura Economica: Buenos Aires.

Rasmussen, S., & Schoen, D. (2010). *Mad as hell: How the tea party movement is fundamentally remaking our two-party system*. New York: Harper.

Riker, W. H. (1980). Implications from the disequilibrium of majority rule for the study of institutions. *American Political Science Review*, *74*, 432–446.

Riker, W. H. (1982). *Liberalism against populism*. San Francisco: Freeman.

Riker, W. H. (1986). *The art of political manipulation*. New Haven: Yale University Press.

Rogowski, R. (1989). *Commerce and coalitions*. Princeton: Princeton University Press.

Sandel, M. J. (2012). *What money can't buy*. New York: Ferrar, Strauss and Giroux.

Schofield, N. (1978). Instability of simple dynamic games. *Review of Economic Studies*, *45*, 575–594.

Schofield, N. (2006a). Equilibria in the spatial stochastic model with party activists. *Review of Economic Design*, *10*, 183–203.

Schofield, N. (2006b). *Architects of political change: Constitutional quandaries and social choice theory*. Cambridge: Cambridge University Press.

Schofield, N. (2007). The mean voter theorem: Necessary and sufficient conditions for convergent equilibrium. *Review of Economic Studies*, *74*, 965–980.

Schofield, N. (2011). *Social choice: Equilibrium or catastrophe*. Working paper: Washington University in St. Louis.

Schofield, N., Miller, G., & Martin, A. (2003). Critical elections and political realignment in the U.S.: 1860–2000. *Political Studies*, *51*, 217–240.

Schofield, N., & Sened, I. (2006). *Multiparty democracy: Elections and legislative politics*. Cambridge: Cambridge University Press.

Schofield, N., & Cataife, G. (2007). A model of political competition with activists applied to elections in Argentina, 1985–1995. *Mathematical Social Sciences*, *53*, 213–231.

Schofield, N., & Miller, G. (2007). Activists and political coalitions in the U.S. *The American Political Science Review*, *51*, 518–531.

Schofield, N., & Levinson, M. (2008). The collapse of authoritarian regimes. *Philosophy, Politics and Economics*, *7*, 243–283.

Schofield, N., Claassen, C., Gallego, M., & Ozdemir, U., (2011a). Empirical and formal models of the United States presidential elections in 2000 and 2004. In N. Schofield & G. Caballero (Eds.), *The Political Economy of Institutions, Democracy and Voting*. Heidelberg: Springer.

Schofield, N., Claassen, C., Ozdemir, U., & Zakharov, A. V. (2011b). Estimating the effects of activists in two-party and multi-party systems: A comparison of the United States in 2008 and Israel in 1996. *Socical Ch Welfare*, *36*, 483–518.

Schofield, N., Gallego, M., & Jeon, J. S. (2011c). Leaders, voters and activists in elections in the Great Britain 2005 and 2010. *Electoral Studies*, *30*, 484–496.

Schofield, N., & Schnidman, E. (2011). Support for political leaders. *Homo Oeconomicus*, *28*, 7–47.

Skocpol, T., & Williamson, V. (2010). *The tea party and the re making of republican conservatism*. Oxford: Oxford: University Press.

Sokoloff, K. L., & Engerman, S. L. (2000). Institutions, factor endowments and the paths of development in the new world. *Journal of Economic Perspectives*, *14*, 217–232.

Stokes, D. (1963). Spatial models and party competition. *American Political Science Review*, *57*, 368–377.

Stokes, D. (1992). Valence politics. In D. Kavanagh (Ed.), *Electoral politics*. Oxford: Clarendon Press

Stokes, S. C. (2005). Perverse accountability: A formal model of machine politics with evidence from Argentina. *American Political Science Review*, *99*, 315–325.

Taussig, F. W. (1888). *The history of the present tariff, 1860–1883*. New York: Putnam.

Tomasky, M. (2010). The money fighting health care reform. *New York Review of Books*, *57*(6), 10–14.

Train, K. (2003). *Discrete choice methods for simulation*. Cambridge: Cambridge University Press.

Wittman, D. (1977). Candidates with policy preferences: A dynamic model. *Journal of Economic Theory*, *14*, 180–189.

Wood, G. S. (2009). *The empire of liberty*. Oxford: Oxford University Press.

Part III
The Measurement of Power

A Review of Some Recent Results on Power Indices

J. M. Alonso-Meijide, B. Casas-Méndez and M. G. Fiestras-Janeiro

1 Introduction

In principle, any democratic institution is designed as systems of governance that reflect the influential effect of each member on any decision making process. Thus, it can be seen as a decision support system in which voters choose among different alternatives. In this chapter we consider the case of binary decision rules in a double sense: each voter chooses between two alternatives; and, once the decision making process ends, voters are divided into two groups (agents are in favor of or against the proposal). Such situations can be represented by simple games where coalitions are classified in two groups: winning or losing in accordance with the success or failure of passing a proposal when all members of a coalition are involved in the decision making process.

The Shapley-Shubik index (Shapley and Shubik 1954) and the non-normalized Banzhaf index (Banzhaf 1965) are the most popular measures in this context. Other measures have been proposed such as the Deegan-Packel index (Deegan and Packel 1978) and the Public Good Index (Holler 1982). While the Shapley-Shubik index and the non-normalized Banzhaf index take into account all coalitions, the

---

This chapter is a reissue based on the work of José M. Alonso-Meijide, Balbina Casas-Méndez and M. Gloria Fiestras-Janeiro (2011), A review of some recent results on power indices, *Essays in Honor of Hannu Nurmi, a special number of Homo Oeconomicus* 28 (1/2): 143–159

J. M. Alonso-Meijide (✉) · B. Casas-Méndez
Department of Statistics and Operations Research, University of Santiago de Compostela, Santiago de Compostela, Spain
e-mail: josemaria.alonso@usc.es

B. Casas-Méndez
e-mail: balbina.casas.mendez@usc.es

M. G. Fiestras-Janeiro
Department of Statistics and Operations Research, University of Vigo, Vigo, Spain
e-mail: fiestras@uvigo.es

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*, DOI: 10.1007/978-3-642-35929-3_12, © Springer-Verlag Berlin Heidelberg 2013

Deegan-Packel and the Public Good Index take only into account the minimal winning coalitions. In the literature there is no consensus about which power index is the most suitable. The study of the mathematical properties of a power index highlights similarities and differences in relation to other indices. Laruelle (1999) provides a good survey on the topic of properties of power indices. In this chapter, we provide some characterizations of the mentioned indices. There are some reasons for getting characterizations of the power indices: first, for a mathematically elegant and pleasant spirit; second, because a set of basic (and assumed independent and hence minimal) properties is a tool to decide on the use of the index. Finally, such a set allows researchers to compare a given value with others and to select the most suitable one for a given problem.

In this chapter we present some results on the study of the Deegan-Packel and the Public Good Index that the research group SaGaTh[1] (Galicia, Spain) has obtained and some results on a new power index, the Shift Power Index, proposed in Alonso-Meijide and Freixas (2010). This research group has also worked in other areas related with power indices. Results and discussion will be presented in Sect. 6. The chapter is organized as follows. Section 2 is devoted to recall some basic definitions on simple games and the study of power indices in order to make the chapter self-contained. Section 3 presents characterizations of the Shapley-Shubik, non-normalized Banzhaf, Deegan-Packel, and Public Good indices. Section 4 is devoted to the Shift Power Index, a new measure based on a desirability relation. In Sect. 5, we propose extensions of the Deegan-Packel and Public Good Index by taking into consideration winning coalitions without null agents.

2 Preliminaries

2.1 Simple Games

A **simple game** is a pair (N, W) where N is a coalition and W is a family of subsets of N satisfying:

(i) $N \in W$, $\emptyset \notin W$ and
(ii) the monotonicity property, i.e.,

$$S \subseteq T \subseteq N \text{ and } S \in W \text{ implies } T \in W.$$

[1] In this research program more colleagues were involved. We specially want to thank our coauthors Mikel Álvarez-Mozos, Mario Bilbao, Carlos Bowles, Francesc Carreras, Joseph Freixas, Julio R. Fernández, Flavio Ferreira, Manfred J. Holler, Silvia Lorenzo-Freire, Stefan Napel, Andreas Nohn, Guillermo Owen, and Alberto Pinto for the discussions and contributions to this research topic. More information about members and working areas of the group is available in the website http://eio.usc.es/pub/io/xogos/.

This representation of simple games follows the approach by Felsenthal and Machover (1998) and by Peleg and Sudhölter (2003). In a simple game, N is the set of members of a committee and W is the set of winning coalitions.

In a simple game (N, W), a coalition $S \subseteq N$ is **winning** if $S \in W$ and is **losing** if $S \notin W$. We denote by $SI(N)$ the set of simple games with player set N.[2]

A winning coalition $S \in W$ is a **minimal winning coalition** (MWC) if every proper subset of S is a losing coalition, that is, $S \subseteq N$ is a MWC in (N, W) if $S \in W$ and $T \notin W$ for any $T \subset S$. We denote by M^W the set of MWC of the simple game (N, W). Given a player $i \in N$ we denote by M_i^W the set of MWC such that i belongs to, that is, $M_i^W = \{S \in M^W / i \in S\}$. Two simple games (N, W) and (N, V) are **mergeable** if for all pair of coalitions $S \in M^W$ and $T \in M^V$, it holds that $S \nsubseteq T$ and $T \nsubseteq S$.

Given two simple games (N, W) and (N, V), we define the union game $(N, W \vee V)$ in such a way that for all $S \subseteq N, S \in W \vee V$ if $S \in W$ or $S \in V$, and we define the intersection game $(N, W \wedge V)$ in such a way that for all $S \subseteq N, S \in W \wedge V$ if $S \in W$ and $S \in V$. If two simple games (N, W) and (N, V) are mergeable, the set of minimal winning coalitions in game $(N, W \vee V)$ is precisely the union of the set of minimal winning coalitions of games (N, W) and (N, V).

A **null player** in a simple game (N, W) is a player $i \in N$ such that $i \notin S$ for all $S \in M^W$. Two players $i, j \in N$ are **symmetric** in a simple game (N, W) if $S \cup i \in W$ if and only if $S \cup j \in W$ for all $S \subseteq N \backslash \{i, j\}$ such that $S \notin W$. Given a player $i \in N$, the **set of swings** of i in the simple game (N, W), $\eta_i(W)$, is formed by the coalitions $S \subseteq N \backslash i$ such that $S \notin W$ and $S \cup i \in W$.

A simple game(N, W) is a weighted majority game if there is a set of weights $w_1, w_2, \ldots, w_n$ for players, with $w_i \geq 0, 1 \leq i \leq n$, and a quota $q \in \mathbb{R}^+$ such that $S \in W$ if and only if $w(S) \geq q$, where $w(S) = \sum_{i \in S} w_i$. A weighted majority game is represented by $[q; w_1, w_2, \ldots, w_n]$. An example of a weighted majority game is provided below.

Example 1 Consider the weighted majority game [14; 6, 5, 4, 4, 3], whose minimal winning coalitions are:

$$M^W = \{\{1, 2, 3\}, \{1, 2, 4\}, \{1, 2, 5\}, \{1, 3, 4\}, \{2, 3, 4, 5\}\}.$$

Coalition $\{1, 2, 3, 4\}$ is a winning coalition, but is not a minimal winning coalition. Coalition $\{1, 2\}$ is a losing coalition. Players 3 and 4 are symmetric and there is no null player.

[2] We will use shorthand notation and write $S \cup i$ for the set $S \cup \{i\}$ and $S \backslash i$ for the set $S \backslash \{i\}$. We will denote by s the number of members in a finite set S.

2.2 Power Indices

A **power index** is a function f which assigns an n-dimensional real vector $f(N, W)$ to a simple game (N, W), where the i-th component of this vector, $f_i(N, W)$, is the power of player i in the game (N, W) according to f. The power index of a simple game can be interpreted as a measure of the ability of the different players to turn a losing coalition into a winning one. In what follows we discuss the Shapley-Shubik index (Shapley and Shubik 1954), the non-normalized Banzhaf index (Banzhaf 1965; Owen 1975), the Deegan-Packel index (Deegan and Packel 1978), and the Public Good Index (Holler 1982). The Shapley-Shubik index depends on the number of permutations on the set of players and of the size of each coalition in each swing. The non-normalized Banzhaf index considers that power depends only on the number of swings of each player and is not directly associated with the order of players. The Deegan-Packel index considers that only minimal winning coalitions will emerge victorious, each minimal winning coalition has an equal probability of forming, and, players in a minimal winning coalition divide the "spoils" equally among the members of this coalition. In a similar way, for the Public Good Index, only minimal winning coalitions are considered relevant when it comes to measuring power. Given a simple game, the Public Good Index of a player is equal to the total number of minimal winning coalitions containing this player divided by the sum of these numbers for all players.

Next, we provide the formal expressions of each one of these indices. Take a simple game (N, W).

I1. The Shapley-Shubik index (SSI) assigns to each player $i \in N$ the real number

$$\varphi_i(N, W) = \sum_{S \in \eta_i(W)} \frac{s!(n-s-1)!}{n!}.$$

I2. The non-normalized Banzhaf index (BI) assigns to each player $i \in N$ the real number

$$\beta_i(N, W) = \frac{|\eta_i(W)|}{2^{n-1}}.$$

I3. The Deegan-Packel index (DP) assigns to each player $i \in N$ the real number

$$\rho_i(N, W) = \frac{1}{|M^W|} \sum_{S \in M_i^W} \frac{1}{|S|}.$$

I4. The Public Good Index (PGI) assigns to each player $i \in N$ the real number

$$\delta_i(N, W) = \frac{\left|M_i^W\right|}{\sum_{j \in N} \left|M_j^W\right|}.$$

Example 2 Using the weighted majority game defined in Example 1, we illustrate the computation of these power indices.

The set of swings of player 1 in this game is:

$$\eta_1(W) = \{\{2,3\},\{2,4\},\{2,5\},\{3,4\},\{2,3,4\},\{2,3,5\},\{2,4,5\},\{3,4,5\}\}.$$

Once the set of swings of each player is obtained, we compute the Shapley-Shubik index and the non-normalized Banzhaf index:

$$\varphi(N,W) = \left(\frac{20}{60},\frac{15}{60},\frac{10}{60},\frac{10}{60},\frac{5}{60}\right) \quad \text{and} \quad \beta(N,W) = \left(\frac{8}{16},\frac{6}{16},\frac{4}{16},\frac{4}{16},\frac{2}{16}\right).$$

The set of MWC such that player 1 belongs to is:

$$M_1^W = \{\{1,2,3\},\{1,2,4\},\{1,2,5\},\{1,3,4\}\}.$$

After taking into account the sets M_i^W for every player $i = 1,2,3,4,5$, we obtain the Deegan-Packel index and Public Good Index:

$$\rho(N,W) = \left(\frac{16}{60},\frac{15}{60},\frac{11}{60},\frac{11}{60},\frac{7}{60}\right) \quad \text{and} \quad \delta(N,W) = \left(\frac{4}{16},\frac{4}{16},\frac{3}{16},\frac{3}{16},\frac{2}{16}\right).$$

3 Some Characterizations

For any power index, understood as a solution concept for simple games, it is always interesting, in both theory and practice, to have the above explicit formulae. Besides, to have a list of properties of an index is desirable with respect to the theoretical discussion and the application of the measures to real-world. Here, we review some of these properties and applied them to SSI, BI, DP, and PGI.

A1. A power index f satisfies **efficiency** if $\sum_{i\in N} f_i(N,W) = 1$, for every simple game (N,W).

A2. A power index f satisfies the **null player** property if $f_i(N,W) = 0$, for every simple game (N,W) and for every null player $i \in N$.

A3. A power index f satisfies **symmetry** if $f_i(N,W) = f_j(N,W)$, for every simple game (N,W), and for every symmetric players $i,j \in N$.

A4. A power index f satisfies the **strong monotonicity** property if $f_i(N,W) \geq f_i(N,V)$, for every simple games (N,W) and (N,V), and for every player $i \in N$ such that $\eta_i(V) \subseteq \eta_i(W)$.

A5. A power index f satisfies the **transfer** property if $f(N,W \vee V) + f(N,W \wedge V) = f(N,W) + f(N,V)$, for every simple games (N,W) and (N,V).

A6. A power index f satisfies the **total power** property if $\sum_{i\in N} f_i(N,W) = \overline{\eta}(W)/2^{n-1}$, for every simple game (N,W), where $\overline{\eta}(W) = \sum_{i\in N}|\eta_i(W)|$.

A7. A power index f satisfies the **DP-mergeability** property if

$$f(N, W \vee V) = \frac{|M^W|f(N,W) + |M^V|f(N,V)}{|M^{W \vee V}|},$$

for every pair of mergeable simple games (N, W) and (N, V).

A8. A power index f satisfies the **PGI-mergeability** property if

$$f(N, W \vee V) = \frac{f(N,W)\sum_{i \in N}|M_i^W| + f(N,V)\sum_{i \in N}|M_i^V|}{\sum_{i \in N}|M_i^{W \vee V}|},$$

for every pair of mergeable simple games (N, W) and (N, V).

A9. A power index f satisfies the **DP-minimal monotonicity** property if

$$|M^W|f_i(N,W) \geq |M^V|f_i(N,V),$$

for every pair of simple games (N, W) and (N, V), and for every player $i \in N$ such that $M_i^V \subseteq M_i^W$.

A10. A power index f satisfies the **PGI-minimal monotonicity** property if

$$f_i(N,W)\sum_{j \in N}\left|M_j^W\right| \geq f_i(N,V)\sum_{j \in N}\left|M_j^V\right|,$$

for every pair of simple games (N, W) and (N, V), and for every player $i \in N$ such that $M_i^V \subseteq M_i^W$.

Property A1 states that the power sharing of the players is equal to 1. Following property A2, a null player gets nothing. Property A3 says that if two players play the same role, then they should have the same power. Property A4 was proposed by Young (1985), and it states that if a player "has more swings" in (N, W) than in (N, V), then, his power cannot be greater in the second game. We want to emphasize that each one of these first four properties use a single game in their definition. Property A5 was proposed by Dubey (1975) as a substitute of Shapley's additivity property that makes no sense in the context of simple games. It establishes that the sum of the power of a player in two simple games coincides with the sum of the power of this player in the union and the intersection game. Property A6 was proposed by Dubey and Shapley (1979). It states that the power of players adds up to the total number of swings divided by the number of coalitions that player i can join. Properties A7 and A8 state that power in the union game is a weighted mean of the power of the two component games, where the weights are provided by taking into account the set of minimal winning coalitions. These two properties provide a link between the power of a player in the union game and the power of the player in the original games. That is, Properties A7 and A8 use three games in their definition.

The last two properties were presented in Alonso-Meijide et al. (2008) and Lorenzo-Freire et al. (2007). In this case, exactly two games are used in the definitions of the properties. These two properties do not establish two equalities, only two inequalities. In the formulation of DP-minimal monotonicity, a relation between a power index in two simple games (N, W) and (N, V) is given in terms of the cardinality of their sets of minimal winning coalitions. This property states that if every minimal winning coalition containing a player $i \in N$ in the game (N, V) is a minimal winning coalition in the game (N, W), then the power of this player in the game (N, W) is not less than his power in the game (N, V) (previously, we must weight this power by the number of minimal winning coalitions of games (N, W) and (N, V), respectively). Note that DP-minimal monotonicity implies

$$f_i(N, W)\left|M^W\right| = f_i(N, V)\left|M^V\right|,$$

for any two simple games (N, W) and (N, V), and for all $i \in N$ such that $M_i^W = M_i^V$. Similarly, the PGI-minimal monotonicity takes into account the relation between a power index in two simple games (N, W) and (N, V). For this property, the weights are given by the number of minimal winning coalitions of every player of games (N, W) and (N, V). The weight $\left|M_j^W\right|$ coincides with the number of swings for a player j in the game (N, W), if we consider minimal winning coalitions only. PGI-minimal monotonicity property implies

$$f_i(N, W)\sum_{j\in N}\left|M_j^W\right| = f_i(N, V)\sum_{j\in N}\left|M_j^V\right|,$$

for any two simple games (N, W) and (N, V), and for all $i \in N$ such that $M_i^W = M_i^V$. In Alonso-Meijide and Holler (2009) this property is discussed in detail.

To end this section, we present some characterizations that enable us to directly compare the power indices I1–I4.

C1. (Dubey 1975) The unique power index f that satisfies transfer, null player, symmetry, and efficiency is the Shapley-Shubik index.
C2. (Dubey and Shapley 1979) The unique power index f that satisfies transfer, null player, symmetry, and total power is the non-normalized Banzhaf index.
C3. (Deegan and Packel 1978) The unique power index f that satisfies DP-mergeability, null player, symmetry, and efficiency is the Deegan-Packel index.
C4. (Holler and Packel 1983) The unique power index f that satisfies PGI-mergeability, null player, symmetry, and efficiency is the Public Good Index.
C5. (Lorenzo-Freire et al. 2007) The unique power index f that satisfies DP-minimal monotonicity, null player, symmetry, and efficiency is the Deegan-Packel index.
C6. (Alonso-Meijide et al. 2008) The unique power index f that satisfies PGI-minimal monotonicity, null player, symmetry, and efficiency is the Public Good Index.

The null player property and symmetry appear in all characterizations presented above. Except the non-normalized Banzhaf index, that satisfies the properties of total power, all indices that we consider satisfy efficiency. In addition to these three properties, namely, symmetry, null player and efficiency (or total power in the case of the non-normalized Banzhaf index), that only use one game in their definitions, another property is necessary to characterize each of these indices. This last property must be one of the group of properties that established an inequality between two games (monotonicity properties) or an equality between the union game and the component games (transfer or mergeability properties). Depending on the property used, each one of the four indices is characterized.

4 The Shift Power Index

Riker (1962, p. 100) claimed that "parties seek to increase votes only up to the size of a minimum coalition". This follows from the well–known "Size Principle" (Riker 1962, p. 32) which implies that, given a multi–member voting body or weighted majority game $[q; w_1, \ldots, w_n]$ a coalition S_0 will be formed provided that $w(S_0) = \min_{S \in W} w(S)$. Taking this principle into account, we could consider that not all the minimal winning coalitions will be formed. In Alonso-Meijide and Freixas (2010) a new power index is proposed. It contains elements of both the Public Good Index and Riker's principle; thus, to certain extent, it can be seen as an intermingled solution. The fundamental idea is the notion of desirability which we recall here.

Let (N, W) be a simple game, i and j be two voters. The *desirability relation* denoted by $\succsim$ is defined in N as follows. We say that i is at least as desirable as j (as a coalitional partner), denoted by $i \succsim j$ if $S \cup \{j\} \in W \;\Rightarrow\; S \cup \{i\} \in W$, for every coalition $S \subseteq N \setminus \{i,j\}$. Player i is said to be (strictly) more desirable than j, denoted by $i \succ j$ if $i \succsim j$ and there is a coalition $T \subseteq N \setminus \{i,j\}$ such that $T \cup \{i\} \in W$ and $T \cup \{j\} \notin W$. Players i and j are said to be equally desirable, denoted by $i \sim j$ if $i \succsim j$ and $j \succsim i$. It is not difficult to see that the desirability relation $(\succsim)$ is a pre-ordering. We say that a simple game (N, W) is **complete or linear** if the desirability relation is a complete pre-ordering. In particular, every weighted majority game is also a complete game, since $w_i \geq w_j$ implies $i \succsim j$. If voters act selfishly looking for the formation of a coalition with partners who are as weak as possible, then under the desirability relation the weaker voters tend to be the most crucial voters at the expense of the stronger ones. The following relative power measure intends to capture this idea.

The proposed index is to a certain extent similar to the Public Good Index and the Deegan-Packel index because it does not take into account surplus coalitions. If S is an arbitrary losing coalition in a simple game (N, W), and $i, j \in N \setminus S$ with $i \succ j$, then i joining S offers more chances to convert S into a winning coalition than j joining S. Precisely, if S is a losing coalition and $S \cup j$ wins, then $S \cup i$ also wins. Thus, if we regard all losing coalitions as a whole, we can assert that i is globally better positioned than j to be chosen as a coalitional partner. Roughly speaking the proposed power index is based on the following principles: (i) every voter wishes to form part of a minimal winning coalition and, (ii) every voter wishes to form part only of those minimal winning coalitions in which no player can be replaced by a weaker one and still winning.

These principles lead us to take a look at the notion of the shift minimal winning coalition which is at the core of the proposed power index. In practice, this notion implies that, for each player, an index will take into account the number of shift minimal winning coalitions that player belongs to, independently of the size of such coalitions.

Let (N, W) be a simple game and $\succsim$ be its desirability ordering. A coalition $S \in M^W$ is **shift minimal** if for every $i \in S$ and $j \notin S$ such that $i \succ j$ it holds $(S \setminus i) \cup j \notin W$. The set of shift minimal winning coalitions will be denoted by SM^W. SM_i^W is the set of shift minimal winning coalitions such that i belongs to, that is, $SM_i^W = \{S \in SM^W / i \in S\}$. Take a simple game (N, W).

I5. The Shift Power Index assigns to each player $i \in N$ the real number

$$\sigma_i(N, W) = \frac{\left|SM_i^W\right|}{\sum_{j=1}^{n}\left|SM_j^W\right|}.$$

The computation of the Shift Power Index involves a subset of the set of coalitions used in the computation of the Public Good Index, because each shift minimal winning coalition is a minimal winning coalition.

Example 3 Using the weighted majority game defined in Example 1 we illustrate the computation of the Shift Power Index. Notice that the complete desirability relation is $1 \succ 2 \succ 3 \sim 4 \succ 5$. Coalition $\{1, 2, 3\}$ is a minimal winning one, but is not shift minimal since $3 \succ 5$ and coalition $\{1, 2, 5\}$ still win. It is not difficult to check that the set of shift minimal coalitions is:

$$SM^W = \{\{1, 2, 5\}, \{1, 3, 4\}, \{2, 3, 4, 5\}\}.$$

The Shift Power Index is:

$$\sigma_i(N, W) = \left(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}\right).$$

We also introduce a property with the flavour of monotonicity written in terms of the shift minimal winning coalitions.

A11. A power index f satisfies the **Shift-minimal monotonicity** property if

$$f_i(N,W)\sum_{j\in N}\left|SM_j^W\right| \geq f_i(N,V)\sum_{j\in N}\left|SM_j^V\right|,$$

for every pair of simple games (N, W) and (N, V), and for every player $i \in N$ such that $SM_i^V \subseteq SM_i^W$.

Using Property A11, and others defined previously, we obtain a characterization of the Shift Power Index.

C7. (Alonso-Meijide and Freixas 2010) The unique power index f that satisfies Shift-minimal monotonicity, null player, symmetry, and efficiency is the Shift Power Index.

5 Quasi-Minimal Coalition Indices

In Alonso-Meijide et al. (2011a), we follow a different point of view. Our intention was to consider as a measure of a player's power two indices similar to those proposed by Deegan and Packel (1978) and Holler (1982) using a set of coalitions bigger than the set of minimal winning coalitions and smaller than the set of winning coalitions. We are interested in that the new indices satisfy the null player property. Then, we define a new family of coalitions.

A winning coalition $S \subseteq N$ is a **quasi-minimal winning coalition** if there is not a null player $i \in S$. The set of quasi-minimal winning coalitions will be denoted by QM^W and the set of quasi-minimal winning coalitions such that i belongs to by QM_i^W, that is, $QM_i^W = \{S \in QM^W / i \in S\}$. It is clear that, for every simple game (N, W),

$$M^W \subseteq QM^W \subseteq W.$$

Example 4 Consider the simple game of four voters, $N = \{1, 2, 3, 4\}$, defined by its set of minimal winning coalitions:

$$M^W = \{\{2,3\},\{2,4\},\{3,4\}\}.$$

Player 1 is the unique null player of the game and coalitions $\{1, 2, 3\}$, $\{1, 2, 4\}$, $\{1, 3, 4\}$, and $\{1, 2, 3, 4\}$ are winning coalitions. However, the set of quasi-minimal winning coalitions is:

$$QM^W = M^W \cup \{\{2,3,4\}\},$$

since $\{2, 3, 4\}$ is also a winning coalition and voter 1 does not belong to it.

We consider a modification of the Deegan-Packel index and a modification of the Public Good Index, defined as follows. Take a simple game (N, W).

I6. The modified Deegan-Packel index assigns to each player $i \in N$ the real number

$$\rho'_i(N, W) = \frac{1}{|QM^W|} \sum_{S \in QM_i^W} \frac{1}{|S|}.$$

I7. The modified Public Good Index assigns to each player $i \in N$ the real number

$$\delta'_i(N, W) = \frac{|QM_i^W|}{\sum_{j \in N} |QM_j^W|}.$$

Álvarez-Mozos (2012) proposes a characterization of each one of these two indices.

Example 5 Once again we use Example 1 to illustrate the computation of these power indices. The set of quasi-minimal winning coalitions such that player 1 belongs to is:

$$QM_1^W = \{\{1,2,3\},\{1,2,4\},\{1,2,5\},\{1,3,4\},$$

$$\{1,2,3,4\},\{1,2,3,5\},\{1,2,4,5\},\{1,3,4,5\},\{1,2,3,4,5\}\}.$$

In this game, the set of null players is empty, thus $QM^W = W$.

After taking into account the sets of quasi-minimal winning coalitions of each one of the rest of players, we obtain the modified Deegan-Packel index and the modified Public Good Index:

$$\rho'(N, W) = \left(\frac{152}{600}, \frac{132}{600}, \frac{112}{600}, \frac{112}{600}, \frac{92}{600}\right) \quad \text{and} \quad \delta'(N, W) = \left(\frac{9}{37}, \frac{8}{37}, \frac{7}{37}, \frac{7}{37}, \frac{6}{37}\right).$$

6 Other Results in Related Research

Our research group has also worked in other areas related with power indices. Next, we briefly summarize these results.

In simple games, a member is considered critical when his elimination from a winning coalition turns this coalition into a losing coalition. Johnston (1978) argued that the non-normalized Banzhaf index, which is based on the idea of a removal of a critical voter from a winning coalition, does not take into account the total number of critical members in each coalition. Clearly, if a voter is the only critical agent in a coalition, this gives a stronger sign of power than in the case where all agents are critical. This is the focal idea underlying the Johnston index. In Lorenzo-Freire et al. (2007), we provide the first characterization of the Johnston index in the class of simple games, using a transfer property that involves the solution for unanimity games of certain coalitions.

For representing social decision situations adequately, sophisticated models have been developed. One of them is the simple game endowed with a priori unions, that is, a partition of the player set which describes a pre-defined (exogenously given) coalition structure. The traditional power indices are not suitable for measuring the distribution of power in these situations because adequate measures of power should take the coalition structure into account. The so-called Owen value (Owen 1977) is an early extension of the Shapley-Shubik index to social decisions with a priori unions. Afterwards, several extensions of the Shapley-Shubik index and the non-normalized Banzhaf index have been proposed for the model of simple games with a priori unions. Alonso-Meijide et al. (2007) compare these different extensions and give arguments to defend the use of them that will depend on the context where they are to be applied. Alonso-Meijide et al. (2009c) introduce adaptations of the conventional monotonicity notions that are suitable for voting games with a priori unions. In Alonso-Meijide et al. (2011b) the Deegan-Packel index is extended to simple games with a priori unions and axiomatically characterized. Of course, this extension coincides with the Deegan-Packel index when each union is formed by only one player or there is only one union. Furthermore, this extension satisfies two types of symmetry, one among players of the same union, and the second one among unions in the game played among unions, and two properties with the same flavour as DP-mergeability.

Two similar versions of mergeability (PGI-mergeability) are quintessential to the two PGI extensions to a priori unions introduced in Alonso-Meijide et al. (2010a). The first one stresses the public good property which suggests that all members of a winning coalition derive equal power, irrespective of their possibility to form alternative coalitions. The second extension follows earlier work on the integration of a priori unions (see Owen 1977). It refers to essential subsets when allocating power shares, taking the outside options of the coalition members into consideration. In Alonso-Meijide et al. (2010b), axiomatizations for six variants of the Public Good Index for games with a priori unions are provided. Two such coalitional PGIs have been introduced and alternatively axiomatized in Alonso-Meijide et al. (2010a). The other four coalitional PGIs have been introduced in Holler and Nohn (2009). The first variant elaborates the original idea of the PGI that the coalitional value is a public good and only minimal winning coalitions of the so called quotient game played by the unions are relevant. The remaining three variants use a two-step distribution where, on the member stage, they take into account the possibilities of players to threaten their partners through leaving their union.

Alonso-Meijide and Casas-Méndez (2007) defined and characterized a modification of the Public Good Index for situations in which some players are incompatible, that is, some players cannot cooperate among them by political reasons. The incompatibilities among players are modelized by a graph in such a way that if two players are linked by an arc of the graph, these two players are incompatible. Alonso-Meijide et al. (2009a) propose a modification of the non-normalized Banzhaf value for this kind of situations, provide two characterizations of it and illustrate it with a real world example taken from the political field.

One of the main difficulties with power indices is that computation generally requires the sum of a very large number of terms. Owen (1972) defined the multilinear extension of a game. Owen (1975) proposed a procedure to compute the Shapley value and the non-normalized Banzhaf value based on the multilinear extension. The multilinear extension is useful in computing the power of large games such as the Presidential Election Game and the Electoral College Game studied by Owen (1972). The multilinear extension approach has two advantages: thanks to its probabilistic interpretation, the central limit theorem of probability can be applied, and, further, it is applied to composition of games. Alonso-Meijide et al. (2008) introduce procedures to calculate the Deegan-Packel index, the Public Good Index and the Johnston index by means of the multilinear extensions.

The generating functions are also efficient tools to compute power indices of weighted voting games. One of the strengths of these procedures is that they allow to get exact values of the indices, even in case we have a very large number of non-null voters. Furthermore, the time required for the computation is in practice much lower than the one required for the computation of the traditional Banzhaf and Shapley-Shubik indices. This method, drawn from Cantor's early work (see Lucas (1983)) and Brams and Affuso (1976), provides algorithms to compute the non-normalized Banzhaf and Shapley-Shubik power indices of weighted voting games. In Alonso-Meijide and Bowles (2005), a procedure based on generating functions is provided to compute power indices in weighted majority games restricted by a priori unions. The method is illustrated by an application to the International Monetary Fund. Alonso-Meijide et al. (2009b) apply the generating functions procedures to the distribution of power in the enlarged European Union, modelled as a weighted multiple majority game restricted by a priori unions.

Acknowledgments We wish to thank the suggestions and comments of the referees. This research received financial support from Ministerio Español de Ciencia y Tecnología through Project ECO2008-03484-C02-02 and from Xunta de Galicia through Project INCITE09-207-064-PR.

References

Alonso-Meijide, J. M., & Bowles, C. (2005). Generating functions for coalitional power indices: An application to the IMF. *Annals of Operations Research*, *137*, 21–44.

Alonso-Meijide, J. M., & Casas-Méndez, B. (2007). The public good index when some voters are incompatible. *Homo Oeconomicus*, *24*, 449–468.

Alonso-Meijide, J. M., & Freixas, J. (2010). A new power index based on minimal winning coalitions without any surplus. *Decision Support Systems*, *49*, 70–76.

Alonso-Meijide, J. M., & Holler, M. J. (2009). Freedom of choice and weighted monotonicity of power. *Metroeconomica*, *60*, 571–583.

Alonso-Meijide, J. M., Carreras, F., Fiestras-Janeiro, M. G., & Owen, G. (2007). A comparative axiomatic characterization of the Banzhaf-Owen coalitional value. *Decision Support Systems*, *43*, 701–712.

Alonso-Meijide, J. M., Casas-Méndez, B., Holler, M. J., & Lorenzo-Freire, S. (2008). Computing power indices: Multilinear extensions and new characterizations. *European Journal of Operational Research*, *188*, 540–554.

Alonso-Meijide, J. M., Álvarez-Mozos, M., & Fiestras-Janeiro, M. G. (2009a). The Banzhaf value when some players are incompatible, *essays in Honor of Hannu Nurmi*, Vol. 1. *Homo Oeconomicus*, *26*, 403–415.

Alonso-Meijide, J. M., Bilbao, J. M., Casas-Méndez, B., & Fernández, J. R. (2009b). Weighted multiple majority games with unions: Generating functions and applications to the European Union. *European Journal of Operational Research*, *198*, 530–544.

Alonso-Meijide, J. M., Bowles, C., Holler, M. J., & Napel, S. (2009c). Monotonicity of power in games with a priori unions. *Theory and Decision*, *66*, 17–37.

Alonso-Meijide, J. M., Casas-Méndez, B., Fiestras-Janeiro, M. G., & Holler, M. J. (2010a). Two variations of the public good index for simple games with a priori unions. *Control and Cybernetics*, *39*, 839–855.

Alonso-Meijide, J. M., Casas-Méndez, B., Fiestras-Janeiro, M. G., Holler, M. J., & Nohn, A. (2010b). Axiomatization of public good indices with a priori unions. *Social Choice and Welfare*, *35*, 517–533.

Alonso-Meijide, J. M., Álvarez-Mozos, M., Ferreira, F., & Pinto, A. (2011a). Two new power indices based on winning coalitions. *Journal of Difference Equations and Applications*, *17*, 1095–1100.

Alonso-Meijide, J. M., Casas-Méndez, B., Fiestras-Janeiro, M. G., & Holler, M. J. (2011b). The Deegan-Packel index for simple games with a priori unions. *Quality and Quantity*, *45*, 425–439.

Álvarez-Mozos, M. (2012). Essays on Cooperative Games with Restricted Cooperative and Simple Games. Ph.D. Thesis. University of Santiago de Compostela: Spain.

Banzhaf, J. F. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review*, *19*, 317–343.

Brams, S. J., & Affuso, P. J. (1976). Power and size: A new paradox. *Theory and Decision*, *7*, 29–56.

Deegan, J., & Packel, E. W. (1978). A new index of power for simple n-person games. *International Journal of Game Theory*, *7*, 113–123.

Dubey, P. (1975). On the uniqueness of the Shapley value. *International Journal of Game Theory*, *4*, 131–139.

Dubey, P., & Shapley, L. S. (1979). Mathematical properties of the Banzhaf power index. *Mathematics of Operations Research*, *4*, 99–131.

Felsenthal, D. S., & Machover, M. (1998). *The measurement of voting power: Theory and practice, problems and paradoxes*. Cheltenham: Edward Elgar.

Holler, M. J. (1982). Forming coalitions and measuring voting power. *Political Studies*, *30*, 262–271.

Holler, M. J., & Packel, E. W. (1983). Power, luck and the right index. *Journal of Economics*, *43*, 21–29.

Holler, M. J., & Nohn, A. (2009). The public good index with threats in a priori unions, *essays in Honor of Harnu Nurmi*, Vol. 1. *Homo Oeconomicus*, *26*, 393–401.

Johnston, R. J. (1978). On the measurement of power: Some reactions to Laver. *Environment and Planning A*, *10*, 907–914.

Laruelle, A. (1999). On the choice of a power index, Instituto Valenciano de Investigaciones Económicas WP-AD 99–10.

Lorenzo-Freire, S., Alonso-Meijide, J. M., Casas-Méndez, B., & Fiestras-Janeiro, M. G. (2007). Characterizations of the Deegan-Packel and Johnston power indices. *European Journal of Operational Research*, *177*, 431–444.

Lucas, W. F. (1983). Measuring power in weighted voting systems. In S.J. Brams, W. Lucas, P.D. Straffin (Eds.), *Political and related models* (pp. 183–238). Springer-Verlag: New York.

Owen, G. (1972). Multilinear extensions of games. *Management Science*, *18*, 64–79.

Owen, G. (1975). Multilinear extensions and the Banzhaf value. *Naval Research Logistics Quarterly*, *22*, 741–750.

Owen, G. (1977). Values of games with a priori unions. In R. Henn, O. Moeschlin (Eds.), *Essays in mathematical economics and game theory* (pp. 76–88). New York: Springer-Verlag.

Peleg, B., Sudhölter, P. (2003). *Introduction to the theory of cooperative games*. Boston: Kluwer Academic Publisher.

Riker, W. H. (1962). *The theory of political coalitions*. New Haven: Yale University Press.

Shapley, L. S., & Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review*, *48*, 787–792.

Young, J. M. (1985). Monotonic solutions of cooperative games. *International Journal of Game Theory*, *14*, 65–72.

Power, Cooperation Indices and Coalition Structures

Rafel Amer and Francesc Carreras

1 Introduction

The main purpose of this chapter is to define a coalition value for TU-games endowed with a cooperation index and a coalition structure. The notion of cooperation index (equivalent to that of weighted hypergraph) was introduced in Amer and Carreras (1995). It provides the foundations for a *quantitative* theory of restricted cooperation that exhibits high precision and flexibility and generalizes several earlier qualitative methods. (For further details on the significance and scope of cooperation indices, we refer the reader to the above reference.) The value we associate with situations described by a game and a cooperation index is a generalization of the Shapley value.

A further step is then suggested by the usefulness of the coalition value as a tool for the analysis of the game dynamics—coalition formation—, which demands an extension of this concept to the new situations we are considering. In fact, this is a crucial point for a full development of the cooperation index theory, because it is only natural to suppose that the greatest incidence of a cooperation index will be precisely found in the bargaining process that leads to the formation of coalitions.

We have tried to find an axiomatic system as simple and powerful as possible to characterize the (generalized) coalition value. A strong symmetry principle (that of *balanced contributions*), already suggested in Myerson (1980) and in Hart and

Earlier versions of this article appeared in 2000 in *Homo Oeconomicus* 17, 11–29, and in 2001 as a chapter of Power Indices and Coalition Formation (M. J. Holler and G. Owen, eds), Kluwer, 153–173.

R. Amer · F. Carreras (✉)
Department of Applied Mathematics II and Industrial and Aeronautical Engineering School of Terrassa, Technical University of Catalonia, Colom 11, 08222 Terrassa, Spain
e-mail: francesc.carreras@upc.edu

R. Amer
e-mail: rafel.amer@upc.edu

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*, DOI: 10.1007/978-3-642-35929-3_13,

Mas-Colell (1989), has been used to this end. As a first application of this principle, we present in Sect. 2 a new axiomatization of the generalized Shapley value for games with a cooperation index. In Sect. 3, the (classical) coalition value is characterized by using two forms of the strong symmetry principle, one for players within any block and another for blocks of the coalition structure, together with a weak version of efficiency.

Next, still in Sect. 3, we define a (generalized) coalition value for game situations where not only a coalition structure but also a cooperation index are given and characterize it, using again two forms of the strong symmetry principle. Besides, some special cases are considered where our new coalition value reduces to more familiar values—in particular to the generalized Shapley value, and not only when the coalition structure is trivial—or remains unaltered after local modifications of the coalition structure.

Finally, two numerical examples are considered in Sect. 4. A concluding remark is in order. Both the Shapley and the coalition value have been commonly used as measures of power by applying them to simple games. One could then ask why we do not reduce ourselves to consider only this kind of games. The answer is that, as will be seen below, modifying a simple game by means of a cooperation index usually produces a non-simple game—from which we derive the generalized Shapley and coalition values—, still having a natural interpretation as a "political game". Hence we develop our theory within the more general framework of (TU) cooperative games, although all our examples will start with a simple game.

1.1 Notation

We shall be concerned with *games* with transferable utility (TU-games), i.e. pairs (N, v) where N is a finite set of *players* and $v : 2^N \longrightarrow \mathbb{R}$ is the *characteristic function*, which assigns to every coalition $S \subseteq N$ a real number $v(S)$ and satisfies $v(\phi) = 0$. A *carrier* for a game (N, v) is a subset $K \subseteq N$ such that $v(S) = v(S \cap K)$ for any $S \subseteq N$. As pointed out in Roth (1988), a player $i \in N$ is *null* in (N, v) if $v(S) = v(S \setminus \{i\})$ for all $S \subseteq N$. The function v is said to be *superadditive* if $v(S \cup T) \geq v(S) + v(T)$ whenever $S \cap T = \phi$. If $T \subseteq S$ we shall write $S_T = S \setminus T$. Given a game (N, v) and a coalition $T \subseteq N$, v_T will denote the restriction of v to 2^{N_T}; it defines a game (N_T, v_T).

Let N be a finite set and let g^N be the set of all unordered pairs (called *links* and written $i : j$) of distinct elements of N. Every $g \subseteq g^N$ is a *graph* on N, and one can then speak of *paths* and *connected components* in any $S \subseteq N$. A *coalition structure* in N is a partition $\mathcal{B}$ of N into nonempty subsets, called *blocks*. If $T \subseteq N$, $\mathcal{B}_T$ will denote the coalition structure induced by $\mathcal{B}$ in N_T. If $I \in \mathcal{B}$, $\mathcal{B}_I$ will denote the coalition structure $\mathcal{B} \setminus \{I\}$ in N_I. When $T = \{i\}$ we shall simply write N_i, v_i and $\mathcal{B}_i$. Finally, given natural numbers $s \leq n$ we define $\gamma(n, s) = \frac{(s-1)!(n-s)!}{n!}$. Further notation will be introduced below.

2 Game Situations and Allocation Rules

To solve a cooperative game is commonly interpreted as defining one or more payoff vectors that might be accepted by all the players according to some rationality criteria. In the case of simple games, which are very often used to represent political decision-making bodies, payoff vectors are usually interpreted as distributions of power among the agents. The Shapley value (Shapley 1953) is an essential contribution to this problem, because it applies to any game (N, v) and selects for it a unique payoff vector $\Phi(N, v) \in \mathbb{R}^N$, which assigns to every player $i \in N$ a payoff denoted by $\Phi_i(N, v)$.

When considering games with additional information—not stored in the characteristic function—the preceding notions need to be generalized. A *game situation* will be a triple $(N, v, \mathcal{I})$, where (N, v) is a game and $\mathcal{I}$ is a mathematical object that contains the external data we wish to take into account to study the game. An *allocation rule* will be a map X that assigns to every game situation $(N, v, \mathcal{I})$ a payoff vector $X(N, v, \mathcal{I}) \in \mathbb{R}^N$.

For instance, the Shapley value itself is an allocation rule for situations where $\mathcal{I} = \phi$; the Myerson value (Myerson 1977) is an allocation rule for situations of the form (N, v, g), where g is a communication graph on N; the Aumann-Drèze value (Aumann and Drèze 1974) and the coalition value (Owen 1977; see also Owen 1995) are allocation rules for situations $(N, v, \mathcal{B})$, where $\mathcal{B}$ is a coalition structure in N.

This section is devoted to recall the generical principle of strong symmetry and to obtain a new characterization for the generalized Shapley value associated with game situations defined by a cooperation index. Roughly speaking, the principle states: *the variation that the payoff to player i undergoes when player j* (*and all additional information concerning him*) *leaves the game must be equal to the variation of the payoff to player j if i leaves the game*.

A version of this principle has already been used by Myerson (1980), with the name of "balanced contributions", to characterize an extension of the Myerson value to NTU-games endowed with a (unweighted) communication hypergraph. Hart and Mas-Colell (1989) use another form of the principle, together with efficiency, to axiomatize the Shapley value. As they point out (about their version), "the principle is a straightforward generalization of the equal division of surplus idea for two-person problems and seems to be a most natural way to compare the relative position (or strengths) of the players."

Here, we wish only to point out that "elementary" proofs (by induction on the number of players) for both the characterization of the Shapley value (without using potential theory) and that of the Myerson value (for TU-games) can be achieved using the standard technique of Theorem 2.1.

For the sake of completeness and also for their repeated use in Sect. 3, we recall some definitions and results from Amer and Carreras (1995). A *game situation with a cooperation index* is a triple (N, v, p) whose third component is a function $p : 2^N \longrightarrow [0, 1]$ such that $p(\{i\}) = 1$ for all $i \in N$. Generalizing Myerson's (1980) terminology, every $T \subseteq N$ such that $p(T) > 0$ will be called a *p-conference*. Overlapping conferences define in a natural way "paths" between players, and hence a notion of connectedness; the connected components, called *islands*, form a partition N/p of N. More generally, $P^+(S,p)$ will denote the set of partitions of a given $S \subseteq N$ into conferences. If $T \subseteq N$, p_T will denote the restriction of p to 2^{N_T}, and will be written p_i when $T = \{i\}$.

Finally, we recall (Amer and Carreras 1995, Theorem 4.1) that there exists a unique allocation rule Ψ, applicable to every game situation with a cooperation index (N, v, p), that satisfies the following axioms:

1. Local superefficiency: for any island $I \in N/p$

$$\sum_{i \in I} \Psi_i(N, v, p) = \max_{P \in P^+(I,p)} \sum_{T \in P} v(T)p(T).$$

2. Fairness: Given $R \subseteq N$ and indices p_1, p_2 such that $p_1(S) = p_2(S)$ for all $S \neq R$,

$$\Psi_i(N, v, p_1) - \Psi_i(N, v, p_2) = \Psi_j(N, v, p_1) - \Psi_j(N, v, p_2) \qquad \forall i, j \in R.$$

This allocation rule, that will be called the *generalized Shapley value*, is defined by $\Psi(N, v, p) = \Phi(N, v/p)$, where $(N, v/p)$ is the p-restricted game given by

$$(v/p)(S) = \max_{P \in P^+(S,p)} \sum_{T \in P} v(T)p(T) \qquad \forall S \subseteq N.$$

Our alternative characterization for Ψ is as follows.

Theorem 2.1 *There exists a unique allocation rule X, applicable to every game situation with a cooperation index (N, v, p), that satisfies the following axioms*:

1. *Local superefficiency*;
2. *Strong symmetry*: *if $i, j \in T$ and T is a p-conference*,

$$X_i(N, v, p) - X_i(N_j, v_j, p_j) = X_j(N, v, p) - X_j(N_i, v_i, p_i).$$

This rule is $X = \Psi$, the generalized Shapley value.

Proof (Existence) It suffices to show that Ψ satisfies strong symmetry. From the fact that $P^+(S, p_i) = P^+(S, p)$ for any $S \subseteq N_i$ it follows that $v_i/p_i = (v/p)_i$ and, given that $\Psi(N, v, p) = \Phi(N, v/p)$, strong symmetry for Ψ is a consequence of the Shapley value strong symmetry (see e.g. Hart and Mas-Colell 1989).

(Uniqueness) Let X^1, X^2 be allocation rules satisfying (1) and (2). We shall show, by induction on $n = |N|$, that $X^1(N, v, p) = X^2(N, v, p)$ for any game situation (N, v, p). If $n = 1$ only local superefficiency matters. Let $n > 1$. For any $i, j \in T$, T being any island, strong symmetry says that

$$X_i^1(N,v,p) - X_i^1(N_j,v_j,p_j) = X_j^1(N,v,p) - X_j^1(N_i,v_i,p_i),$$
$$X_i^2(N,v,p) - X_i^2(N_j,v_j,p_j) = X_j^2(N,v,p) - X_j^2(N_i,v_i,p_i).$$

Subtracting these equalities and using the inductive hypothesis it follows that

$$X_i^1(N,v,p) - X_i^2(N,v,p) = X_j^1(N,v,p) - X_j^2(N,v,p),$$

and hence the function d, given by $d(i) = X_i^1(N,v,p) - X_i^2(N,v,p)$, is a constant function on each island $I \in N/p$. Using now local superefficiency, $X_i^1(N,v,p) = X_i^2(N,v,p)$ is true within each island, and hence in N. □

Let us consider, now, a first example of application of the generalized Shapley value that shows how to get an a priori evaluation of the power distribution when a cooperation index matters.

Example 2.2 Let (N,v) be the straight majority game $[2;1,1,1]$, that is, the game where

$$v(S) = 1 \quad \text{if} \quad |S| \geq 2, \qquad v(S) = 0 \quad \text{otherwise},$$

and let p be the cooperation index given as follows:

$$p(\{i\}) = 1 \quad \text{for} \quad i = 1,2,3; \quad p(\{1,2\}) = 0.7;$$
$$p(\{1,3\}) = 0.5; \quad p(\{2,3\}) = 0; \quad p(N) = 0.$$

The modified game $(N, v/p)$ is given by

$$(v/p)(\{i\}) = 0 \quad \text{for} \quad i = 1,2,3; \quad (v/p)(\{1,2\}) = 0.7;$$
$$(v/p)(\{1,3\}) = 0.5; \quad (v/p)(\{2,3\}) = 0; \quad (v/p)(N) = 0.7$$

and the modified Shapley value is

$$\Psi(N,v,p) = \Phi(N,v/p) = (0.4333, 0.1833, 0.0833),$$

whereas the Shapley value for the original game is constant:

$$\Phi(N,v) = (0.3333, 0.3333, 0.3333).$$

The modified value reflects that player 1 is the best placed to form coalitions and to take profit; player 2 is also in a better position than player 3. In the original game the players shared 1, whereas in the modified one this amount is reduced to 0.7. A possible interpretation of this fact is the following: if the players play the game many times, they will form different coalitions depending on the play, but, in the end, they will share an average of 0.7 per play (if the cooperation index remains unaltered). This seems to be the case of political parties that, instead of forming a coalition for the entire legislature, sign partial commitments with different groups depending on the issue and obtain, then, an inefficient distribution of power.

3 Coalition Value and Cooperation Indices

The coalition value is an allocation rule for game situations with a coalition structure $(N, v, \mathcal{B})$. It differs from the Aumann-Drèze value in that it does not consider $\mathcal{B}$ as a final structure for the game but as a starting point for further negotiations at a higher level (that of blocks in the quotient game): this difference is reflected in the efficiency condition. The coalition value generalizes the Shapley value, which arises not only when $\mathcal{B} = \{N\}$ but also when $\mathcal{B} = \{\{i\}/i \in N\}$: these are the so-called trivial structures. Characterizations somewhat different from the original one may be found, e.g., in Hart and Kurz (1983) and Winter (1992).

Our first result in this section states a new axiomatization for the coalition value, using the strong symmetry principle at two levels: for individuals (players) and for blocks. Note that this allows us to dispense with the null-player and additivity axioms, which are necessary in other formulations mentioned above. It will be useful, throughout this section, to write

$$X_K(N, v, \mathcal{I}) = \sum_{i \in K} X_i(N, v, \mathcal{I})$$

for any allocation rule X and any $K \subseteq N$.

Theorem 3.1 *There exists a unique allocation rule X, applicable to every game situation with a coalition structure $(N, v, \mathcal{B})$, that satisfies the following axioms:*

1. *Efficiency*: $\sum_{i \in N} X_i(N, v, \mathcal{B}) = v(N)$;
2. *Block strong symmetry: for all $I, J \in \mathcal{B}$*

$$X_I(N, v, \mathcal{B}) - X_I(N_J, v_J, \mathcal{B}_J) = X_J(N, v, \mathcal{B}) - X_J(N_I, v_I, \mathcal{B}_I);$$

3. *Inner strong symmetry: for all $K \in \mathcal{B}$ and all $i, j \in K$*

$$X_i(N, v, \mathcal{B}) - X_i(N_j, v_j, \mathcal{B}_j) = X_j(N, v, \mathcal{B}) - X_j(N_i, v_i, \mathcal{B}_i).$$

This rule is $X = \hat{y}$, the coalition value.

Proof (Existence) Since our efficiency axiom already appears in other axiomatizations of the coalition value (e.g. Winter 1992), it is sufficient to check that $\hat{y}$ satisfies both strong symmetry postulates. Let $\mathcal{B} = \{K^1, K^2, \ldots, K^m\}$, $M = \{1, 2, \ldots, m\}$ be a set of representatives and (M, u) be the quotient game of (N, v) by $\mathcal{B}$, defined by

$$u(T) = v\left(\bigcup_{q \in T} K^q\right) \qquad \forall T \subseteq M.$$

(Block strong symmetry) We know (Owen 1977) that for any block, say, $K = K^r$,

$$\hat{y}_K(N, v, \mathcal{B}) = \sum_{i \in K} \hat{y}_i(N, v, \mathcal{B}) = \Phi_r(M, u).$$

Moreover, if $I = K^i$ and (M_i, w) is the quotient game of (N_I, v_I) by $\mathcal{B}_I$, it follows that $w = u_i$. The Shapley value strong symmetry yields therefore, for all blocks $I = K^i$ and $J = K^j$,

$$\begin{aligned}\hat{y}_I(N, v, \mathcal{B}) - \hat{y}_I(N_J, v_J, \mathcal{B}_J) &= \Phi_i(M, u) - \Phi_i(M_j, u_j) \\ &= \Phi_j(M, u) - \Phi_j(M_i, u_i) = \hat{y}_J(N, v, \mathcal{B}) - \hat{y}_J(N_I, v_I, \mathcal{B}_I).\end{aligned}$$

(Inner strong symmetry) Let $K = K^r \in \mathcal{B}$ and assume that $i, j \in K$. The explicit formula for the coalition value (Owen 1977) gives

$$\hat{y}_i(N, v, \mathcal{B}) = \sum_{\substack{T \subseteq M \\ r \notin T}} \sum_{\substack{S \subseteq K \\ i \in S}} \gamma(m, t+1)\gamma(k, s)[v(\tilde{T} \cup S) - v(\tilde{T} \cup S_i)]$$

where $\tilde{T} = \cup_{q \in T} K^q$, and m, t, k and s are, respectively, the cardinalities of M, T, K and S. Applying again Owen's formula we obtain

$$\hat{y}_i(N_j, v_j, \mathcal{B}_j) = \sum_{\substack{T \subseteq M \\ r \notin T}} \sum_{\substack{S \subseteq K_j \\ i \in S}} \gamma(m, t+1)\gamma(k-1, s)[v(\tilde{T} \cup S) - v(\tilde{T} \cup S_i)].$$

A straightforward calculation leads to the following equality:

$$\begin{aligned}&\hat{y}_i(N, v, \mathcal{B}) - \hat{y}_i(N_j, v_j, \mathcal{B}_j) \\ &\quad= \sum_{\substack{T \subseteq M \\ r \notin T}} \sum_{\substack{S \subseteq K \\ i,j \in S}} \gamma(m, t+1)\gamma(k, s)[v(\tilde{T} \cup S) - v(\tilde{T} \cup S_i) - v(\tilde{T} \cup S_j) + v(\tilde{T} \cup S_{\{i,j\}})].\end{aligned}$$

The symmetrical appearance of i and j in this expression justifies that

$$\hat{y}_i(N, v, \mathcal{B}) - \hat{y}_i(N_j, v_j, \mathcal{B}_j) = \hat{y}_j(N, v, \mathcal{B}) - \hat{y}_j(N_i, v_i, \mathcal{B}_i).$$

(Uniqueness) Let X^1, X^2 be allocation rules satisfying (1)–(3). We shall show, by induction on $n = |N|$, that $X^1(N, v, \mathcal{B}) = X^2(N, v, \mathcal{B})$ for all game situations $(N, v, \mathcal{B})$. For $n = 1$ this follows from efficiency. Let $n > 1$. Block strong symmetry and the inductive hypothesis imply that the function d, defined by $d(I) = X^1_I(N, v, \mathcal{B}) - X^2_I(N, v, \mathcal{B})$ for every $I \in \mathcal{B}$, is constant. Efficiency implies that d vanishes. Inner strong symmetry and the induction hypothesis apply to prove that the function δ_I, defined by $\delta_I(i) = X^1_i(N, v, \mathcal{B}) - X^2_i(N, v, \mathcal{B})$ within each block $I \in \mathcal{B}$, is constant on I. In fact, δ_I vanishes too, since $0 = d(I) = \sum_{i \in I} \delta_I(i)$, and hence $X^1_i(N, v, \mathcal{B}) = X^2_i(N, v, \mathcal{B})$ for all $i \in I$ and all $I \in \mathcal{B}$, i.e. for all $i \in N$. $\square$

The coalition value has been used to study the dynamics of coalition formation in game situations of types that could be described by particular cooperation indices. For example, in Carreras and Owen (1988) and (1996), Carreras et al. (1993), and Bergantiños (1993): all these papers show applications of game theory to political science. The procedure is always based on *the computation of the coalition value for the* $\mathcal{I}$-restricted game under different coalition structures that are considered "plausible" according to $\mathcal{I}$. This suggests, in fact, that a set of new allocation rules—one for each type of situation—can be defined in this way, and it seems therefore interesting to find a common axiomatic characterization for them. (We also provide in Sect. 4 two numerical examples illustrating that procedure.)

In order to generalize the coalition value to qualitatively restricted game situations, the case of $(N, v, g, \mathcal{B})$ is perhaps a most basic step in this program and, indeed, what might be called the Myerson coalition value has been defined and axiomatized in Vázquez-Brage et al. (1996). In the present work, we complete this approach by adopting the widest point of view, that of the cooperation indices, and study therefore a game situation of the form $(N, v, p, \mathcal{B})$. A (generalized) coalition value is defined and characterized using the strong symmetry principle.

Let us consider a game situation $(N, v, p, \mathcal{B})$, where p is a cooperation index and $\mathcal{B}$ is a coalition structure. We will use definitions and results from Amer and Carreras (1995) that have been remembered in Sect. 2, just preceding Theorem 2.1. One more definition is needed: we shall say that two $\mathcal{B}$-blocks are linked by p if there exists a p-conference that intersects both blocks. This defines a graph on $\mathcal{B}$ and allows us to speak of paths between blocks: the connected components of $\mathcal{B}$ relatively to this graph will be called *superblocks*.

We define an allocation rule V, applicable to every game situation of the form $(N, v, p, \mathcal{B})$, as follows:

$$V(N, v, p, \mathcal{B}) = \hat{y}(N, v/p, \mathcal{B}),$$

where v/p is the p-restricted game. We call V the *generalized coalition value*.

Theorem 3.2 *There exists a unique allocation rule X, applicable to every game situation $(N, v, p, \mathcal{B})$, that satisfies the following axioms*:

1. *Local superefficiency: for every island $I \in N/p$*

$$\sum_{i \in I} X_i(N, v, p, \mathcal{B}) = \max_{P \in P^+(I,p)} \sum_{T \in P} v(T)p(T);$$

2. *Block strong symmetry: if $R, S \in \mathcal{B}$ are linked by p*

$$\begin{aligned} X_R(N, v, p, \mathcal{B}) &- X_R(N_S, v_S, p_S, \mathcal{B}_S) \\ &= X_S(N, v, p, \mathcal{B}) - X_S(N_R, v_R, p_R, \mathcal{B}_R); \end{aligned}$$

3. *Inner strong symmetry: for all $R \in \mathcal{B}$ and all $i, j \in R$*

$$X_i(N, v, p, \mathcal{B}) - X_i(N_j, v_j, p_j, \mathcal{B}_j) = X_j(N, v, p, \mathcal{B}) - X_j(N_i, v_i, p_i, \mathcal{B}_i).$$

This rule is $X = V$, the generalized coalition value.

Proof (Existence) We will prove that V satisfies properties (1)–(3). Let $(N, v, p, \mathcal{B})$ be a given situation. (Local superefficiency) Define, for any island $I \in N/p$, a game u^I in N as follows:

$$u^I(S) = \max_{P \in P^+(I \cap S, p)} \sum_{T \in P} v(T)p(T) \qquad \forall S \subseteq N.$$

If $N/p = \{I_1, I_2, , I_k\}$, it can be shown (Amer and Carreras 1995, Proposition 3.5) that

$$v/p = \sum_{r=1}^{k} u^{I_r}.$$

Since each $I \in N/p$ is a carrier for u^I, the additivity of the classical coalition value yields, for every island I,

$$\sum_{i \in I} V_i(N, v, p, \mathcal{B}) = \sum_{i \in I} \hat{y}_i(N, u^I, \mathcal{B}) = u^I(N) = \max_{P \in P^+(I, p)} \sum_{T \in P} v(T)p(T).$$

(Block strong symmetry) This directly follows from the block strong symmetry of the coalition value $\hat{y}$, because $v_R/p_R = (v/p)_R$ for every block $R \in \mathcal{B}$. (Inner strong symmetry) The property derives, once more, from the corresponding property of the classical coalition value, using in this case that $v_i/p_i = (v/p)_i$ for any $i \in N$.

(Uniqueness) Let X^1, X^2 be allocation rules that satisfy (1)–(3). We shall show, by induction on $n = |N|$, that $X^1(N, v, p, \mathcal{B}) = X^2(N, v, p, \mathcal{B})$ for any game situation $(N, v, p, \mathcal{B})$. For $n = 1$ use local superefficiency only. Let $n > 1$. Block strong symmetry and the inductive hypothesis imply, as in Theorem 3.1, that the function d, defined by $d(R) = X^1_R(N, v, p, \mathcal{B}) - X^2_R(N, v, p, \mathcal{B})$ for every $R \in \mathcal{B}$, is constant on every superblock of $\mathcal{B}$. Then, every superblock $\mathcal{U}$ is, as a subset of N, isolated with respect to p-connectedness, and hence $\mathcal{U}$ is the union of some islands, say, $I_1, I_2, \ldots, I_p$. If $R_1, R_2, \ldots, R_q$ are the blocks which form $\mathcal{U}$, we have

$$\begin{aligned}
\sum_{h=1}^{q} d(R_h) &= \sum_{h=1}^{q} X^1_{R_h}(N, v, p, \mathcal{B}) - \sum_{h=1}^{q} X^2_{R_h}(N, v, p, \mathcal{B}) \\
&= \sum_{t=1}^{p} X^1_{I_t}(N, v, p, \mathcal{B}) - \sum_{t=1}^{p} X^2_{I_t}(N, v, p, \mathcal{B}) \\
&= \sum_{t=1}^{p} (v/p)(I_t) - \sum_{t=1}^{p} (v/p)(I_t) = 0,
\end{aligned}$$

and, d being a constant function on $\mathcal{U}$, it follows that $d(R) = 0$ for any R in $\mathcal{U}$ and therefore for any $R \in \mathcal{B}$, since $\mathcal{U}$ was arbitrary. Finally, let $i, j \in R \in \mathcal{B}$. Inner strong symmetry and the inductive hypothesis apply to show that the function d_R, defined by $d_R(i) = X^1_i(N, v, p, \mathcal{B}) - X^2_i(N, v, p, \mathcal{B})$ for every $i \in R$, is constant.

Using that $\sum_{i\in R} d_R(i) = d(R) = 0$ we conclude that d_R vanishes. Thus, we have $X_i^1(N,v,p,\mathcal{B}) = X_i^2(N,v,p,\mathcal{B})$ for every $i \in R$ and every $R \in \mathcal{B}$, i.e. for every $i \in N$. □

Now, let us describe the behavior of the generalized coalition value V in some special cases of game situations of the form $(N,v,p,\mathcal{B})$.

Examples 3.3 (a) If $p(S) = 1$ for all $S \subseteq N$, then $v/p = v^e$ (the superadditive extension of v) and $V(N,v,p,\mathcal{B}) = \hat{y}^e(N,v,\mathcal{B})$, where $\hat{y}^e$ denotes the IR-*coalition value* defined by $\hat{y}^e(N,v,\mathcal{B}) = \hat{y}(N,v^e,\mathcal{B})$. If, moreover, v is superadditive, then $V(N,v,p,\mathcal{B}) = \hat{y}(N,v,\mathcal{B})$.

(b) When $\mathcal{B} = \{N\}$, block strong symmetry does not matter, whereas the two other axioms in Theorem 3.2 become those imposed to Ψ in Theorem 2.1; therefore, $V(N,v,p,\{N\}) = \Psi(N,v,p)$.

(c) In a similar way, if $\mathcal{B} = \{\{i\}/i \in N\}$, inner strong symmetry does not say anything, the two other axioms coincide with those of Theorem 2.1 and, again, $V(N,v,p,\{\{i\}/i \in N\}) = \Psi(N,v,p)$.

(d) Assume v is superadditive, let $\mathcal{B}$ be arbitrary and take $p = p_{\mathcal{B}}$, defined by

$$p_{\mathcal{B}}(S) = \begin{cases} 1 & \text{if } S \subseteq K \text{for some } K \in \mathcal{B}, \\ 0 & \text{otherwise.} \end{cases}$$

The meaning of $p_{\mathcal{B}}$ is obvious: the players may freely negotiate among them within each block, but they cannot communicate at all with players belonging to other blocks. Then, $V(N,v,p_{\mathcal{B}},\mathcal{B}) = \Phi_0(N,v,\mathcal{B}) = \Psi(N,v,p_{\mathcal{B}})$, where Φ_0 is the Aumann-Drèze value. Note that the generalized values V and Ψ coincide, even though $\mathcal{B}$ is not trivial but arbitrary, and that the Aumann-Drèze value is shown to be a particular case of the generalized coalition value.

To analyze the following examples we need some elementary properties of the coalition value. Every game v in N can be uniquely written as $v = \sum_{T\subseteq N} a_T u_T$, where u_T is the T-unanimity game in N and $a_T = \sum_{R\subseteq T}(-1)^{t-r}v(R)$ for every nonempty coalition $T \subseteq N$ (t and r are the cardinalities of T and R).

Remark 3.4 If K is a carrier for (N,v), then $a_T = 0$ for all T not contained in K. From this it follows immediately that if a block of $\mathcal{B}$ is a carrier for (N,v) then $\hat{y}(N,v,\mathcal{B}) = \Psi(N,v)$.

Remark 3.5 Let (N,v) be a game and let $\mathcal{B} = \{K_1,K_2,\ldots,K_m\}$ be a coalition structure such that $K_1 = K_0' \cup K_1'$, $K_0' \cap K_1' = \emptyset$ and all members of K_0' are null players in (N,v). Let $\mathcal{B}' = \{K_0',K_1',K_2,\ldots,K_m\}$. Then

$$\hat{y}(N,v,\mathcal{B}) = \hat{y}(N,v,\mathcal{B}').$$

Examples 3.6 (a) If $\mathcal{B} = N/p$, then $V(N,v,p,N/p) = \Psi(N,v,p)$, since applying the coalition value additivity to $v/p = \sum_{I\in N/p} u^I$ (recall the proof of Theorem 3.2) yields $V(N,v,p,N/p) = \sum_{I\in N/p} \hat{y}(N,u^I,N/p)$, and, using that each island $I \in N/p$ is a carrier for (N,u^I), Remark 3.4 above gives

$$V(N,v,p,N/p) = \sum_{I\in N/p} \hat{y}(N,u^I,N/p) = \Psi(N,v/p) = \Psi(N,v,p).$$

We meet again a situation where $\mathcal{B}$ is not trivial but the generalized values V and Ψ coincide.

(b) Let $(N,v,p,\mathcal{B})$ be a game situation where $\mathcal{B} = \{K_1,K_2,\ldots,K_m\}$, and assume that $K_1 = K_0' \cup K_1'$, $K_0' \cap K_1' = \emptyset$ and K_0', K_1' are nonempty and lie in different islands. Let $\mathcal{B}' = \{K_0',K_1',K_2,\ldots,K_m\}$. Therefore

$$V(N,v,p,\mathcal{B}) = V(N,v,p,\mathcal{B}').$$

To prove this we apply again the coalition value additivity to $v/p = \sum_{I\in N/p} u^I$ and use Remark 3.5 to obtain

$$V(N,v,p,\mathcal{B}) = \sum_{I\in N/p} \hat{y}(N,u^I,\mathcal{B}) = \sum_{I\in N/p} \hat{y}(N,u^I,\mathcal{B}') = V(N,v,p,\mathcal{B}').$$

Some comments are in order: we have never demanded any kind of compatibility, between the cooperation index p (or its islands) and the coalition structure $\mathcal{B}$, to formalize the theory in this section. But, as follows from our latter statement, the players will have no interest in forming blocks with members of other islands, and therefore *the only interesting coalition structures are, in practice, those where each island splits into blocks*; thus, superblocks (defined before Theorem 3.2) are reduced to be islands.

(c) Let us assume, now, that a block is included in an island but is not connected (by conferences). Then, this block cannot split into connected components without changing the coalition value, as the following counterexample shows. Let $N = \{1,2,3\}$ and $v(S) = 1$ if $|S| \geq 2$, $v(S) = 0$ otherwise. Let p be the cooperation index defined by $p(S) = 0$ if $S = \{2,3\}$ and $p(S) = 1$ otherwise, and let $\mathcal{B} = \{\{1\},\{2,3\}\}$. Then, $(v/p)(S) = v(S)$ if $S \neq \{2,3\}$ and $(v/p)(\{2,3\}) = 0$; thus, $V(N,v,p,\mathcal{B}) = (0.50,0.25,0.25)$. If block $\{2,3\}$ is subdivided, a new coalition structure $\mathcal{B}' = \{\{1\},\{2\},\{3\}\}$ arises, for which $V(N,v,p,\mathcal{B}') = (0.66,0.16,0.16)$.

An obvious question: if players 2 and 3 cannot communicate because $p(\{2,3\}) = 0$, how can they form a block in $\mathcal{B}$? The null cooperation index assigned to $\{2,3\}$ means that they will not agree to form a coalition, but they may agree in other questions, e.g. in that they will never form separately a coalition with player 1! A situation of this kind arose at the beginning of 1993 in the Parliament of Aragón (Spain). Parties 1 and 2 were holding a coalition government, but a proposal of party 2 about a deep amendment of the Autonomy Statute was refused with the votes of parties 1 and 3 (2 is a regionalist party, whereas 1 and 3 are the main parties at the national level and are not especially inclined to give further competences to regions).

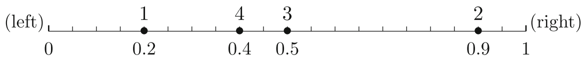


Fig. 1 Party-distribution on a left-to-right axis in Example 4.1

4 Two Examples

This final section is devoted to considering two numerical instances where we shall use the generalized coalition value as a "dynamic" measure of power, leaving to the generalized Shapley value the "static" role of describing the *initial conditions* for the bargaining.

Example 4.1 Let us consider a parliamentary body where four parties share 50 seats, giving rise to the weighted majority game $(N, v) \equiv [26; 20, 15, 11, 4]$. Formally, this is a very simple situation. Any coalition formed by two of the three main parties is *stable*, because the classical coalition value allocates 0.5 units to each one of its members and they cannot better this allocation by going elsewhere. But, perhaps, things are not so simple. Assume that parties are politically located in a classical left-to-right axis as is shown in Fig. 1.

To take into account this ideological component, we introduce a cooperation index p, derived from the distances between parties. For instance, we find

$$p(\{1,2\}) = 1 - d(1,2) = 0.3000$$
$$p(\{1,3,4\}) = 1 - \sqrt{d(1,4)^2 + d(4,3)^2} = 0.7764$$

and so on, until

$$p(N) = 1 - \sqrt{d(1,4)^2 + d(4,3)^2 + d(3,2)^2} = 0.5417.$$

The modified game is therefore

$$\begin{aligned} v/p = {} & 0.3\, u_{\{1,2\}} + 0.7\, u_{\{1,3\}} + 0.6\, u_{\{2,3\}} \\ & - 0.9\, u_{\{1,2,3\}} + 0.1615\, u_{\{1,2,4\}} + 0.0764\, u_{\{1,3,4\}} - 0.1615\, u_N. \end{aligned}$$

From looking at the initial configuration, given by

$$\Psi(N, v, p) = (0.2389, 0.1635, 0.3351, 0.0389),$$

it follows that party 3 is really the strongest player. An evaluation of the generalized coalition value $V(N, v, p, \mathcal{B})$ for different coalition structures—essentially, those where only a v-winning coalition forms—is provided in Table 1, where the first row gives the generalized Shapley value. It tells us that $\mathcal{B} = \{\{1\}, \{2,3\}, \{4\}\}$ is the only stable one, and yields

$$V(N, v, p, \mathcal{B}) = (0.0755, 0.2519, 0.4236, 0.0255).$$

Table 1 Generalized coalition values in Example 4.1

	1	2	3	4
No coalition	0.2389	0.1635	0.3351	0.0389
$\{1,2\}$	0.3139	0.2385	0.1716	0.0524
$\{1,3\}$	0.3210	0.0000	0.4172	0.0382
$\{2,3\}$	0.0755	0.2519	0.4236	0.0255
$\{1,2,3\}$	0.2326	0.1635	0.3422	0.0382
$\{1,2,4\}$	0.3210	0.2519	0.1575	0.0460
$\{1,3,4\}$	0.3139	0.0000	0.4236	0.0389
$\{2,3,4\}$	0.0882	0.2385	0.4172	0.0326

(Note that player 1 is indifferent between coalitions $\{1,3\}$ and $\{1,2,4\}$, player 2 is indifferent between $\{2,3\}$ and $\{1,2,4\}$, player 3 between $\{2,3\}$ and $\{1,3,4\}$, and player 4 would prefer that coalition $\{1,2\}$ forms. Player 4 might thus promote coalition $\{1,2,4\}$ and leave it immediately, but the residual coalition $\{1,2\}$ would then dissolve because player 2 prefers to enter $\{2,3\}$).

Thus, coalition $\{2,3\}$ is not "winning" in the classical sense. It obtains $V_{\{2,3\}}(N,v,p,\mathcal{B}) = 0.6755$, which is more than $(v/p)(\{2,3\}) = 0.6000$ but far from $(v/p)(N) = 0.7764$, and the difference is allocated to players 1 and 4 (which is no longer a dummy player because of its central position). This may be interpreted as caused by disagreements between players 2 and 3, which will be often obliged to negotiate with 1 and/or 4. On the other hand, player 3 receive much more than player 2, and this is in accordance with Owen's (1977) intracoalitional bargaining model if one compares the allocations to players 1, 2, and 3 under $\{1,2\}$ and $\{1,3\}$ in v/p. Summing up, the modified model seems to provide a more realistic view of the political complexity of this situation.

Before proceeding with our second example—the analysis of a real world situation—, the question of how to determine a cooperation index is worthy of mention. The cooperation degree of a given coalition (say, of parties) may depend on many factors: pure ideological positions, strategic conveniences, past experience, future compromises, existence of simultaneous settings where the involved parties (or some of them) are meeting and probably bargaining... In Example 4.1 we have suggested a way for computing the cooperation index exclusively in terms of the left-to-right ideological positions. If one wants to take into account additional components that influence the relationships between parties, two main procedures seem to be plausible.

The first one is purely theoretical, and needs to assume that every factor can be numerically described. In this case, the basic point will be to find a function, of as many variables as factors we have so defined, mapping in a reasonable way the domain of these variables (a cartesian product) into the interval $[0,1]$ of the real line. In our opinion, this is an interesting field of research.

The second method is rather of empirical nature, no necessarily more subjective than the former and, surely, easier to use in practice. By enquiring appropriate people, one can obtain a good estimation of the cooperation degree of every coalition. "Appropriate people" means here, e.g., party leaders or spokesmen, political scientists and observers, mass media (press, radio, television) commentators and, still better, a combination of all of them. A comparison of parties' programatic manifestos should be added as a complementary way (last, but not least) for obtaining the cooperation degree of any coalition. Other possibilities will be welcome.

In the following example we have assumed the role of political observers, and have established what we feel is a reasonable cooperation index in view of the actual behavior of the involved parties.

A final remark on this question. The coalition value is a continuous function of the unanimity coordinates of the game, for it is linear. On the other hand, the unanimity coordinates of the modified game v/p are easily seen to be continuous functions of the cooperation index p, viewed as a vector variable in the $(2^n - n - 1)$-dimensional unit cube. Thus we conclude that the generalized coalition value is a continuous function of the cooperation index, and hence small enough errors in evaluating p will give rise to negligible differences in our analysis of the coalition dynamics of a game by means of V (as can be checked in Example 4.1).

Example 4.2 The case of the *Congreso de los Diputados* (Lower House of the Spanish Parliament) during the 1993–1996 Legislature will be studied here. Eleven parties elected members to the *Congreso* in June 1993, giving rise to the following weighted majority game:

$$(N, v) \equiv [176; 159, 141, 18, 17, 5, 4, 2, 1, 1, 1, 1].$$

This is an "apex game", because all but the four main parties are null players and the set of minimal winning coalitions is

$$W^m = \{\{1,2\}, \{1,3\}, \{1,4\}, \{2,3,4\}\}.$$

Disregarding the null players, we obtain a 4-person game whose Shapley value is given by

$$\Phi(N, v) = (0.5000, 0.1667, 0.1667, 0.1667).$$

Formally, the coalition formation is easy to analyze. Only the minimal winning coalitions are stable, and yield the following coalition values:

$$\hat{y}(N, v, \{1,2\}) = (0.6667, 0.3333, 0, 0)$$

(and analogous results for the two other binary coalitions) and

$$\hat{y}(N, v, \{2,3,4\}) = (0, 0.3333, 0.3333, 0.3333)$$

(an oversized majority such as $\{1, i, j\}$ allocates the same payoff to player 1 as $\{1, i\}$ or $\{1, j\}$, but divides 0.3333 equally among i and j and is therefore not interesting for these two players).

There are, however, complex relationships between the Spanish parties. We shall introduce a cooperation index to describe the political structure and will then obtain very different and much more clear results in analyzing the coalition formation. For a better understanding of our index, let us first identify the agents of our game.

Player 1 is the *Partido Socialista Obrero Español* (PSOE), a left-to-center party that obtained absolute majority in 1982 and 1986 and just 175 seats in 1989. Player 2 is the *Partido Popular* (PP), a strong and growing center-right party. Player 3 is *Izquierda Unida* (IU), a coalition headed by the old communist party. Player 4 is *Convergència i Unió* (CiU), a regionalist middle-of-the-road coalition which was enjoying absolute majority in the Catalonian Parliament since 1984 and was very interested in influencing the national policy without entering the government.

Finally, we shall also mention player 5: it is the *Partido Nacionalista Vasco* (PNV), another middle-of-the-road regionalist party that is holding with PSOE a coalition government in the Basque Country since 1986. The reason to include it is that this party seems to be very important from the cooperation index point of view; it fails, however, to escape from its dummy position when we modify the game.

We shall define a cooperation index that takes into account these political characteristics of the parties. Three reasonable assumptions will make our task easier:

1. Coalitions S of more than 3 parties are highly improbable, and will then be assigned $p(S) = 0$.
2. We need to specify $p(S)$ for winning coalitions only (recall the definition of the modified game v/p).
3. Once $p(S)$ is given for every coalition S such that $|S| = 2$, we assume that, for every T with $|T| = 3$,

$$p(T) = \min\{p(S) : S \subseteq T, |S| = 2\}.$$

These prerequisites and our own opinion about the relationships between parties give rise to a cooperation index that we describe as follows:

$$\begin{array}{lll} p(\{1,2\}) = 0, & p(\{2,4\}) = 0.4, & p(\{1,3,4\}) = 0.2, \\ p(\{1,3\}) = 0.3, & p(\{2,5\}) = 0.5, & p(\{1,3,5\}) = 0.2, \\ p(\{1,4\}) = 0.9, & p(\{3,4\}) = 0.2, & p(\{1,4,5\}) = 0.9, \\ p(\{1,5\}) = 1.0, & p(\{3,5\}) = 0.2, & p(\{2,3,4\}) = 0.1, \\ p(\{2,3\}) = 0.1, & p(\{4,5\}) = 1.0, & \end{array}$$

and $p(S) = 0$ otherwise if $|S| > 1$. The modified game is therefore

$$v/p = 0.3\, u_{\{1,3\}} + 0.9\, u_{\{1,4\}} - 0.3\, u_{\{1,3,4\}} + 0.1\, u_{\{2,3,4\}} - 0.1\, u_{\{1,2,3,4\}}.$$

Table 2 Generalized coalition values in Example 4.2

	1	2	3	4
No coalition	0.4750	0.0083	0.0583	0.3583
$\{1,2\}$	0.4833	0.0167	0.0500	0.3500
$\{1,3\}$	0.5083	0.0000	0.0917	0.3000
$\{1,4\}$	0.5083	0.0000	0.0000	0.3917
$\{1,2,3\}$	0.5083	0.0083	0.0833	0.3000
$\{1,2,4\}$	0.5083	0.0083	0.0000	0.3833
$\{1,3,4\}$	0.4833	0.0000	0.0583	0.3583
$\{2,3,4\}$	0.4000	0.0167	0.0917	0.3917

Notice that, in spite of its very good degree of affinity with two basic parties (PSOE and CiU), PNV is still a null player after modifying the game; thus it will be left aside definitively in our study of the coalition bargaining. Also notice that $(v/p)(N) = 0.9$, and therefore no coalition will share more than this amount among their members. We have already given an interpretation of this inefficiency elsewhere in this chapter. By looking at v-winning coalitions with no more than 3 players, the generalized coalition value gives the results contained in Table 2.

A comparison of the Shapley value for games v and v/p (first row of Table 2) tells us that PP and IU lose power, PSOE remains more or less equally, and CiU gets a much better position. There are three stable coalitions: the first one is $\{1,4\}$, whose players, PSOE and CiU, share all the available power; the second is $\{1,3\}$, formed by PSOE and IU, but they control only a power of 0.6 and leave therefore 0.3 to CiU; the third stable coalition is $\{2,3,4\}$ and, again, its members, PP, IU, and CiU, obtain only 0.5 in all, thus leaving an important fraction of power in PSOE's hands. Finally, note that PSOE holds its optimal value also in oversized coalitions ($\{1,2,3\}$ and $\{1,2,4\}$), but they are not stable.

One can then conclude that any observer of the Spanish political life should agree with this mathematical description of the strategic and ideological tensions at the *Congreso* during the last Legislature. In particular, our result would be found satisfactory because, indeed, a parliamentary coalition between PSOE and CiU has been supporting a minority government of the socialist party. Furthermore, the sharing of power among these two parties, given by the fourth row of Table 2, corresponds very closely to a generalized opinion that PSOE's cabinet has been dominated by the conditional support of CiU, which has very often imposed its criteria on economic and regional (autonomic) policies.

Acknowledgments Research supported by Grant SGR 2009-01029 of the Catalonia Government (*Generalitat de Catalunya*) and Grant MTM 2012-34426 of the Economy and Competitiveness Spanish Ministry.

References

Amer, R., & Carreras, F. (1995). Games and cooperation indices. *International Journal of Game Theory*, *24*, 239–258.

Aumann, R. J., & Drèze, J. (1974). Cooperative games with coalition structure. *International Journal of Game Theory*, *3*, 217–237.

Bergantiños, G. (1993). Aportaciones a la teoría de juegos (in Spanish). Ph. D. Thesis. University of Santiago de Compostela, Spain.

Carreras, F., & Owen, G. (1988). Evaluation of the Catalonian Parliament 1980–1984. *Mathematical Social Sciences*, *15*, 87–92.

Carreras, F., & Owen, G. (1996). An analysis of the Euskarian Parliament. In N. Schofield (Ed.), *Collective decision-making: Social choice and political economy* (pp. 251–263). Boston: Kluwer Academic Publishers.

Carreras, F., García-Jurado, I., & Pacios, M. A. (1993). Estudio coalicional de los parlamentos autonómicos españoles de régimen común (in Spanish). *Revista de Estudios Políticos*, *82*, 159–176.

Hart, S., & Kurz, M. (1983). Endogeneous formation of coalitions. *Econometrica*, *51*, 1047–1064.

Hart, S., & Mas-Colell, A. (1989). Potential, value and consistency. *Econometrica*, *57*, 589–614.

Myerson, R. B. (1977). Graphs and cooperation in games. *Mathematics of Operations Research*, *2*, 225–229.

Myerson, R. B. (1980). Conference structures and fair allocation rules. *International Journal of Game Theory*, *9*, 169–182.

Owen, G. (1977). Values of games with a priori unions. In O. Moeschlin & R. Henn (Eds.), *Essays in mathematical economics and game theory* (pp. 76–88). Berlin: Springer.

Owen, G. (1995). *Game theory* (3rd ed.). San Diego: Academic Press.

Roth, A. E. (Ed.). (1988). *The Shapley value: Essays in honor of L. S. Shapley*. Cambridge: Cambridge University Press.

Shapley, L. S. (1953). A value for n-person games. In A. W. Tucker & H. Kuhn (Eds.), *Contributions to the theory of games II (Annals of Mathematical Studies, 28)* (pp. 307–317). Princeton: Princeton University Press.

Vázquez-Brage, M., García-Jurado, I., & Carreras, F. (1996). The Owen value applied to games with graph-restricted communication. *Games and Economic Behavior*, *12*, 42–53.

Winter, E. (1992). The consistency and potential for values of games with coalition structure. *Games and Economic Behavior*, *4*, 132–144.

The Power of a Spatially Inferior Player

Mika Widgrén and Stefan Napel

1 Introduction

Power is an important concept in the analysis of economic and political institutions, and even of moral codes and ethics. Though everybody has some understanding about who under what circumstances exerts power, the concept is elusive. Therefore, it is not surprising that there is considerable controversy as to what constitutes an appropriate measure of power even in the restricted class of those economic or political institutions which can be represented as simple games in coalitional form.

Power indices assign to each player of a n-person simple game, such as a weighted multi-party voting game, a non-negative real number which indicates the player's a priori power to shape events. Numerous indices have been proposed—most notably by Shapley and Shubik (1954), Banzhaf (1965), Deegan and Packel (1979), and Holler and Packel (1983).[1] On the surface, the distinction between these is whether minimal winning coalitions, crucial coalitions, player permutations, or other concepts are the primitives for measurement. More fundamentally,

This article was first published in *Homo Oeconomicus* 19(3), 327–343, 2002
Mika Widgrén unexpectedly passed away on 16.8.2009 at the age of 44.

[1] For recent comparative investigations of power indices, their properties and applicability, see Felsenthal and Machover (1998) and Holler and Owen (2001).

M. Widgrén
Turku School of Economics and Public Choice Research Centre,
University of Turku, 20014 Turku, Finland

S. Napel (✉)
Department of Economics, University of Bayreuth, Bayreuth 95440, Germany
e-mail: stefan.napel@uni-bayreuth.de

S. Napel
Public Choice Research Centre, University of Turku, Turku, Finland

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_14,

the discussion is about the realism of the distinct probability models behind alternative indices, desired properties like monotonicity and, importantly, the congruence of indications for basic reference cases with predictions by other tools of economic or political analysis.

In this light, the following basic example is striking. Consider the 3-player simple game where the only winning coalitions are the grand coalition ABC and the two coalitions AB and AC. A could be the federal government that needs approval from one of two provincial governments to pass laws. Or, A might be a shareholder who needs to be backed by at least one of two (smaller) shareholders to decide on strategic questions of corporate policy. Economic equilibrium analysis would claim A to be "on the short side of the market", implying that B and C cannot influence terms of trade. From the point of non-cooperative game theory, A can be imagined to make an ultimatum offer to B, asking for approval in return for an only marginal (and in the limit non-existent) concession to B's political or economic interests. A rational player B would have to accept since a potential threat of colluding with C to obtain a better deal is not credible or subgame perfect. A symmetric argument applies to C. Drawing on cooperative game theory, the core and nucleolus of this game are both $\{(1,0,0)\}$ and further support the intuition that B and C are powerless in this game. Despite this clear prediction by different types of non-cooperative and cooperative game-theoretic reasoning, the power indices of Banzhaf and Shapley-Shubik indicate substantial power for powerless players B and C. They yield the power vectors $(\frac{3}{4}, \frac{1}{4}, \frac{1}{4})$ and $(\frac{2}{3}, \frac{1}{6}, \frac{1}{6})$, respectively. In Napel and Widgrén (2001) the notion of *inferior players* was defined to reach a more satisfactory solution.

To put it in a more general context, the criticism that power indices usually face stems from two factors. First, closely related to our example above, traditional power indices do not take players' strategic interaction into account. Second, their capability of modelling complicated institutional features, like agenda-setting, is limited. The inferior player axiom and the strict power index derived from it are an attempt to tackle these problems. In this chapter, we take one further step and attempt to define the concept of inferior players in a spatial voting context. Our goal is then to build an a priori measure of power, which corresponds with the strict power index and which opens the avenue for taking preferences into the analysis of power. Moreover, the approach allows us to model more complex institutional features of the game, like agenda setting.

The fundamental difference between spatial voting and coalitional form games is that the latter has the set of players and the former the set of policy outcomes as the domain. In spatial voting, players are supposed to have ideal points in a policy space and payoff is assumed to be monotonically decreasing in the distance between ideal policy outcome and actual one. In coalitional form games, coalitions rather than individuals gain when a coalition is able to pass proposals. Power indices then give estimates for an individual's influence on a coalition's achievement. In this chapter, we discuss this difference and aim to take both approaches into account.

Recently, strategic aspects and power indices have been studied by Steunenberg et al. (1999). In their analysis, the *strategic power index (StPI)*[2] of player i, Ψ_i, is defined as

$$\Psi_i = \frac{\Delta_d - \Delta_i}{\Delta_d}$$

where Δ_d is the expected distance between the equilibrium outcome and the ideal point of a dummy player, and Δ_i is the expected distance between the equilibrium outcome and the ideal point of player i. In Steunenberg et al. (1999), a dummy is a player who is like an outside observer of the game having no power. The ideal policy of a dummy player is assumed to vary within the same range as the ideal points of the actual players of the game but a dummy does not have any decision making rights and, thus, she does not matter for the outcome of the game.[3] From the formula it is easy to see that a player who always gets her ideal policy obtains one as the power value and a player who is like a dummy gets zero as her power value.

In the case of the strategic power index, the introduction of a dummy player is due only to standardisation. The distance Δ_i plays the key-role. Without normalisation, Δ_i would take the role of an "absolute" power measure. The underlying idea is simply to define power on the basis of proximity between players' most preferred positions and actual outcomes. At first glance this may sound appealing but there is at least one caveat. Given that players' preferences are spatial a voter may well have an ideal point very close to the outcome although the passage of the proposal that has lead to this outcome was completely out of her control. Proximity is often due to luck, not power. Let us illustrate things with a simple example.

Consider a seven-player symmetric perfect information voting game, with player set $\{A, B, C, D, E, F, G\}$ and a 5/7th majority rule. Assume ideal points in a uni-dimensional policy space which order the players' positions from left to right as follows: *ABCDEFG*. Consider a proposal χ which is located in between E's and F's ideal points, but closer to E than F. StPI suggests that if χ is accepted, then E exerts more power in this preference configuration than players A, B, C, D, F and G. However, the outcome of this vote depends on the location of the current state of affairs, i.e. status quo. For simplicity, suppose that status quo lies left of A. Coalition *ABCDE* is then a potential minimal winning coalition, as well as *BCDEF* and *CDEFG*. Consider the first alternative. Given the locational assumptions *ABCDE* cannot be minimal winning with respect to proposal χ in a spatial sense since if the players in it accept proposal χ, then so do F and G. Player A is the most likely member of *ABCDE* to reject the proposal χ but is no longer critical given F's and G's acceptance. We get *BCDEF* as the next candidate minimal winning coalition. The same argument as before holds for this coalition—it is not a minimal

[2] We use this abbreviation instead of SPI to avoid confusion with the *Strict Power Index*, which is defined in Napel and Widgrén (2001) and abbreviated as SPI.

[3] Note that this is not the standard way to define a dummy player.

winning coalition in a spatial sense since if it were approving proposal χ then G would also accept it. Consider *CDEFG*. In this coalition, E is closer to the proposal than any other player. But is this due to her power? The only player who has a credible swing in this coalition is C. Only if status quo is further to the left from C than χ is to the right, C accepts the proposal. Player E does not have such position for this preference configuration. In fact, this holds for nearly all proposals and locations of status quo.

In this chapter, we take an alternative approach to strategic power and follow internal rather than external normalisation. This means that whether a player is dummy or not depends on her capabilities in the game. Contrary to Steunenberg et al. (1999), we assume that any player, dummy or not, is an actual player and not an external observer. We define *a posteriori power* as having an effective pivotal position for a *given* preference configuration, and *(a priori) power* as the ex ante expectation of it, taken with respect to the probabilities of different preference configurations. This allows for different informational considerations and makes the analysis more procedural than in the case of StPI. Our approach leads to a definition of power, which, in fact, corresponds to that of established power indices.

2 Coalitional Form and Spatial Voting Games

Coalitional form voting games deal with all possible coalitions of members of a set $N \equiv \{1, \ldots, n\}$ of players. Players' preferences are not known. Coalitions $S \subseteq N$ are either winning or losing, implying a partition of the set of all coalitions, $\mathcal{P}(N)$, into the set $\mathcal{W}$ of *winning coalitions* and the set $\mathcal{L}$ of *losing coalitions*.[4] A *coalitional form voting game* is a special instance of a *simple game* defined by the pair $(N, \mathcal{W})$, where the set of winning coalitions, $\mathcal{W}$, can be characterized by a non-negative real vector $r_v = (m; w_1, \ldots, w_n)$, where w_i is player i's number of votes and m is the number of votes that establishes a winning coalition. In a *simple majority voting game*, $w_i = 1$ for every player $i \in N$ and $m = n/2 + 1$ or $m = (n+1)/2$ for even or odd n, respectively.

A game $(N, \mathcal{W})$ can equivalently be described by its *characteristic function* v. It maps n-tuples $s \in \{0, 1\}^n$, which represent a feasible coalition $S \subseteq N$ by indicating which players $i \in N$ belong to S ($s_i = 1$) and which do not ($s_i = 0$), either to $v(S) = 1$ if $S \in \mathcal{W}$ or to $v(S) = 0$ if $S \in \mathcal{L}$. When $\mathcal{W}$ represents winning coalitions in a voting game, v is monotonic, i.e. $v(S) = 1$ implies $v(T) = 1$ for any superset $T \supseteq S$.

[4] We only consider *proper games* in which the complement of a winning coalition is losing, i. e. $S \in \mathcal{W} \Rightarrow N - S \in \mathcal{L}$. We do not assume that the game is *decisive*, i.e. additionally $N - S \in \mathcal{L} \Rightarrow S \in \mathcal{W}$, because this would preclude the analysis of qualified majority voting. If both $S \in \mathcal{L}$ and $N - S \in \mathcal{L}$, then status quo prevails (see definition below).

A player who by leaving a winning coalition $S \in \mathcal{W}$ turns it into a losing coalition $S - \{i\} \in \mathcal{L}$ has a *swing* in S. He is called a *crucial* or *critical member* of coalition S. Coalitions in which at least one member is crucial are called crucial coalitions.[5] Coalitions where player i is critical are called *crucial coalitions with respect to i*. Let

$$C_i(v) \equiv \{S \subseteq N \mid v(S) = 1 \wedge v(S - \{i\}) = 0\}$$

denote the set of crucial coalitions w.r.t. i. The number of swings of player i in simple game v is thus

$$\eta_i(v) \equiv \mid C_i(v) \mid .$$

A player i who is never crucial, i.e. $\eta_i(v) = 0$, is called *dummy player*. In Napel and Widgrén (2001), the following related concept is introduced:

Definition 1 Player i is inferior in simple game v if $\exists j \neq$ i:

$$\begin{aligned} & \forall S \in C_i(v): \quad j \in S \\ \wedge \quad & \exists S' \in C_j(v): \quad i \notin S' \end{aligned}$$

An inferior player i is equivalently characterized by $C_i(v) \subsetneq C_j(v)$ for $j \neq i$.[6] It is straightforward to see that every dummy player is inferior but the reverse does not hold (see Napel and Widgrén 2001). Let us refer to a player who is not inferior as *superior*.

The game with $\mathcal{W} = \{AB, AC, ABC\}$ was used above to illustrate the divergence between power predictions based on conventional indices, on the one hand, and competitive analysis or the concept of the core of a game, on the other hand. Imagine that the spoils of a winning coalition in v are \$100 and to be split among its members. Alternatively, consider 100 policy units, e.g. referring to different topics in a policy proposal for each of which the players have distinct preferred alternatives. Regardless of the precise object of conflicting interests in v, player A is in the position of the proposer in a non-cooperative *Ultimatum Game* with B as responder when the situation permits negotiations before the final establishment of a winning coalition. Since A has the option to form a winning coalition without B, B cannot do better but to accept whatever A proposes in terms of B's share of spoils or political influence. A anticipates this and rationally offers B a share of

[5] Deegan and Packel (1979) use the term 'minimal winning coalition', Felsenthal and Machover (1998) the term 'vulnerable coalition' instead of 'crucial coalition'. We, like other authors, follow Bolger's (1980) conceptualization.

[6] Note that the first part of Definition 1 implies that j belongs to all *minimal* winning coalitions to which i belongs and is (by definition) crucial in these. Assuming that i has a swing in a non-minimal winning coalition without j also having one leads to a contradiction once non-crucial members of that coalition—including j—are dropped. Therefore, i does never have a swing without j also having a swing in the same coalition—but j has at least one swing in a coalition without i having one.

(almost) nothing. The Banzhaf index of this game is $\left(\frac{3}{4}, \frac{1}{4}, \frac{1}{4}\right)$. In (Napel and Widgrén 2001) this was corrected by replacing the conventional dummy player axiom of power measurement with a corresponding inferior player axiom. In the example, we get the following strict power index $\left(\frac{3}{4}, 0, 0\right)$.

It is worth noting that in the spatial context considered in this chapter things are different. Despite the fact that the agenda-setter is superior to all voters, the game is *not* necessarily a pure ultimatum game. Basically, this is due to the possible veto power exerted by the voters. The equilibrium outcome of the game depends on the pivotal player's preferred point. This implies that a pivot may be able to put credible threats on the agenda-setter, despite being inferior in a coalitional form, non-spatial sense.

In coalitional form games players' preferences do not have any role in determining the outcome. A usual way to justify this is to say that coalitional form voting games analyse institutions rather than actual votes and that there is no sufficient a priori information about players' preferences. Games in coalitional form thus analyse several votes.

Coalitional form games usually do not model agenda setting either. To add agenda setting into our model let us distinguish between two types of agents, namely fixed agenda setters $j \in A$ and voters $i \in N$. In spatial voting games, players' preferences restrict the class of feasible winning coalitions.

Definition 2 A one-dimensional $(\mathrm{n}+\mathrm{s})$-player spatial voting game with agenda setting is a 5-tuple $(\mathrm{N}, \mathrm{A}, \mathcal{W}, \Lambda, \sigma)$ where N is the set of voters, A is the set of agenda setters, $\mathcal{W}$describes the class of majority coalitions needed for the passage of agenda setters' proposals, $\Lambda = (\lambda_1, \ldots, \lambda_n) \in \boldsymbol{R}^{\mathrm{n}}$ is the vector of voters' ideal points, and $\sigma \in \boldsymbol{R}^{\mathrm{s}}$ is the vector of agenda setters' ideal points.

Throughout this chapter, we assume that A is a singleton, hence $s = 1$. In general, however, it is easy to find examples where the agenda is set by a group of agents using pre-determined rules how to decide upon the agenda. The European Commission serves as an example.

We simplify the model by restricting the analysis to only one policy dimension. We also disregard weighted voting for the sake of simplicity. Given two policies x and y, ideal points partition N into

$$N_{x \succsim y} \equiv \{i \,|\, d(x, \lambda_i) \leq d(y, \lambda_i)\}$$
$$N_{x \prec y} \equiv \{i \,|\, d(x, \lambda_i) > d(y, \lambda_i)\}$$

where $d(a, b) \equiv \sqrt{a^2 + b^2}$ denotes the Euclidian distance between a and b. We normalize the *status quo* to $Q = 0$. Spatial preferences are e.g. represented by the utility functions $\pi_\sigma(\Omega) = -(\sigma - \Omega)^2$ and $\pi_{\lambda_i}(\Omega) = -(\lambda_i - \Omega)^2$, where Ω denotes the policy outcome of the game.

We think of voters' and the agenda setter's ideal points as being *ex ante*—when institutional a priori power is evaluated—random variables denoted by $\tilde{\Lambda} = (\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n)$ and $\tilde{\sigma}$. Their distributions are $F_{\tilde{\Lambda}}$ and $F_{\tilde{\sigma}}$, respectively. However, actual

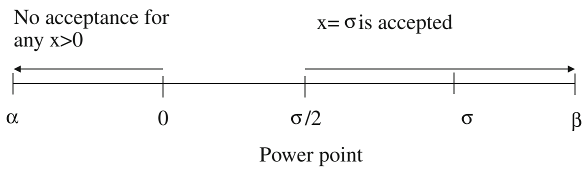


Fig. 1 A simple uni-dimensional policy space

decisions and the pivotal positions that are our indicators of a posteriori power[7] are determined under complete information, i.e. for particular commonly observed realizations Λ and σ of ideal points. For any given preference configuration we consider the following *agenda setting game (ASG)*:

1. Agenda setter A makes a take-it-or-leave-it proposal $\chi = \chi(\Lambda, \sigma, m)$, where m is the number of voters whose acceptance is needed to pass a proposal.
2. Voters $i \in N$ simultaneously accept or reject the proposal. The outcome of the game is $\Omega = \chi$ if the proposal is accepted and $\Omega = 0$ if the proposal is rejected.

A possible policy space of this game is shown in Fig. 1. Suppose that voters' ideal points λ_i are a priori uniformly distributed on $[\alpha, \beta]$ where $\alpha \leq 0, \beta > 0$. Agenda setter's ideal point σ is supposed to lie in $[0, \beta]$ and it is assumed to have uniform distribution. At first glance, this particular assumption may sound restrictive but, in fact, it is with little loss of generality. Assuming $\alpha = 0$ we get the special case where there is no asymmetry with respect to possible ideal points between voters and the agenda setter. Our desire is to generalise the assumption of identical domains for all players' ideal points in a tractable way. For the simple procedural setting above it is natural to concentrate on asymmetry between voters and the agenda setter.

This allows for two kinds of interesting considerations. First, the interval $[\alpha, 0]$ gives the range where a voter does not gain from any proposal made by a rational agenda setter. It is a well-known result from spatial voting that players located in opposite directions from status quo do not cooperate. If we interpret the agenda setter as a seller and voters as buyers, then the interval $[\alpha, 0]$ gives the range where there are no gains from exchange; in a political context, the players have interests so conflicting that no mutually beneficial compromise about changing the status quo is possible. This can also be seen from the individual rationality constraints for acceptance that can be written

$$(\lambda_i - \chi)^2 \leq (\lambda_i - 0)^2 \qquad (IR_i)$$

for voters $i \in N$ and correspondingly

[7] We do not explicitly analyse agenda-setting power here. For a corresponding extension see the general framework in Napel and Widgrén (2004).

$$(\sigma-\chi)^2 \leq (\sigma-0)^2 \qquad (IR_\sigma)$$

for the agenda setter A.

Second, asymmetrically distributed ideal points imply that the status quo bias plays a role in the model. This in turn makes it possible to investigate the effects of inefficiencies on power. Inefficiency emerges when a group of players is able to bloc any proposal made by the agenda setter whose IR_σ-constraint restricts the domain of proposals to $[0, \beta]$. This reduces power of both the agenda setter and those voters who would have preferred to replace status quo by some $\chi > 0$.

Using the assumptions above we get the following cumulative distributions functions

$$F_{\tilde{\lambda}_i}(x) = \begin{cases} 0, & \text{if } x \leq \alpha \\ \frac{x-\alpha}{\beta-\alpha}, & \text{if } \alpha < x \leq \beta \\ 1, & \text{if } x > \beta \end{cases}$$

and

$$F_{\tilde{\sigma}}(x) = \begin{cases} 0, & \text{if } x \leq 0 \\ \frac{x}{\beta}, & \text{if } 0 < x \leq \beta \\ 1, & \text{if } x > \beta. \end{cases}$$

Note that

$$\hat{\lambda}_i \equiv \frac{\tilde{\lambda}_i - \alpha}{\beta - \alpha} \sim U(0, 1).$$

For future use let us define the following re-scaling

$$\Pi(\sigma) \equiv \frac{\frac{1}{2}\sigma - \alpha}{\beta - \alpha}$$

and refer to it as the *power point.* The power point turns out to be the dividing line between cases when a player may exert power and when she may not. The range between the status quo and the power point is crucial for our concept of spatial inferiority. Note that a priori the power point is random. We get

$$\hat{\Pi} \equiv \frac{\frac{1}{2}\tilde{\sigma} - \alpha}{\beta - \alpha} \sim U\left(\frac{-\alpha}{\beta - \alpha}, \frac{\frac{1}{2}\beta - \alpha}{\beta - \alpha}\right).$$

Thinking of the players as representatives for some constituency or organization, it is reasonable to assume in the following that player i votes for a proposal χ whenever (IR_i) is satisfied. This means that after χ is proposed the coalition $N_{\chi \succsim 0} \subseteq N$ will form.

This assumption imposes considerable structure on the coalitions that are formed. Let (i) denote the player j whose ideal point, λ_j, turns out to be the ith smallest of all voters so that $\lambda_{(1)} \leq \ldots \leq \lambda_{(n)}$. The agenda setter's rationality implies $\chi \geq 0$. Thus whenever $(IR_{(k)})$ is satisfied, then so are $(IR_{(k+1)}), \ldots, (IR_{(n)})$. Hence, any coalition which is formed is *convex* or *connected* in the following sense:

$$(i) \in S \wedge (i+l) \in S \Rightarrow (i+1), \ldots, (i+l-1) \in S, \quad l \geq 2.$$

We will refer to a coalition with this property given a realized vector of voter's ideal points, Λ, as a Λ *-connected coalition.*[8]

Whether a particular Λ-connected coalition $S \subseteq N$ will be formed or not depends on the agenda setter's proposal χ. For given χ and Λ, there is a unique (χ, Λ) *-individually rational* or (χ, Λ) *-IR* connected coalition $S = N_{\chi \succsim 0}$ which will form. This may be winning or losing.

In a winning (χ, Λ)-IR coalition S, some players can have a swing in the traditional coalitional sense, i.e. can turn S into a losing coalition by leaving. In a spatial context, threatening to reject χ is generally no credible option e.g. for player (n), who is in fact the most eager to replace the status quo by χ. The swing position which has to be taken seriously by the agenda setter is that of the crucial member of S who is least eager to replace the status quo. With this in mind, we say that player i has a Λ-*spatial swing* in winning coalition S or is Λ-*pivotal* if i has a swing in S and no other player $j \neq i$ with $d(\lambda_j, 0) \leq d(\lambda_i, 0)$, i. e. who is even less eager to replace the status quo by χ, has a swing in S. We call a player $i(\chi, \Lambda)$-*pivotal* in S to abbreviate that i is Λ-pivotal in S and S is (χ, Λ)-IR. In above setting, a winning (χ, Λ)-IR coalition S has to have at least m members, and only player $(n-m+1)$ can have a spatial swing.

To highlight the link between the spatial and the simple coalitional framework, one may define[9]

$$\mathcal{C}_i(\chi, \Lambda) \equiv \begin{cases} \{S\} & \text{if } i \text{ has a } (\chi, \Lambda)\text{-spatial swing in } S \\ \varnothing & \text{otherwise.} \end{cases}$$

By considering all possible (χ, Λ)-combinations one then obtains the set of crucial coalitions with respect to i defined above, i.,e.

$$\bigcup_{\substack{\chi \in [0, \beta], \\ \Lambda \in [\alpha, \beta]^n}} \mathcal{C}_i(\chi, \Lambda) = \mathcal{C}_i(v).$$

The refinement of swings to spatial swings captures one criterion for a crucial position to mean power in a decision framework with explicit spatial preferences. However, for a player to be truly powerful, his preferences should *matter* in terms

---

8 In our setting, one might use the more specific term Λ-*right* -connected coalition to stress that a formed coalition necessarily includes all players to the right of a given member.

9 Note that $S = N_{\chi \succsim 0}$ is the only (χ, Λ)-IR coalition, meaning that $\mathcal{C}_i(\chi, \Lambda)$ is well-defined.

of outcome, i.e. a small change of preferences should lead to a small change of outcome. This requires a spatial swing, but having one is not sufficient. Consider the 7-player game above and assume, for instance, $\lambda_A < \lambda_B < 0 < \lambda_C < \ldots < \lambda_G$,[10] i.e. only players C to G may prefer to replace the current state of affairs by some proposal $\chi \geq 0$. For $\sigma > 0$, the agenda setter wants to replace the status quo. The χ closest to his ideal point σ which establishes a (χ, Λ)-IR winning coalition S is his optimal proposal χ^*. For $0 < \sigma < \lambda_C$, $\chi^*(\sigma, \Lambda) = \sigma$ is the optimal proposal, and will become the policy outcome of the game. Player C's spatial swing position does *not* have any effect on the outcome in this case. In fact, C's preferences do not influence the outcome until $\sigma > 2\lambda_C$ holds.

Given the assumptions made for above agenda setting game, we get the following subgame perfect Nash equilibrium proposal[11]

$$\chi^*(\sigma, \Lambda) = \chi^*\left(\sigma, \lambda_{(n-m+1)}\right) = \begin{cases} \sigma & \text{if } \lambda_{(n-m+1)} \geq \frac{1}{2}\sigma \\ 2\lambda_{(n-m+1)} & \text{if } \lambda_{(n-m+1)} \in (0, \frac{1}{2}\sigma) \\ 0 & \text{if } \lambda_{(n-m+1)} \leq 0 \end{cases}$$

which is accepted by voters $(n), \ldots, (n-m+1)$ and by any $(n-m), \ldots, (l)$, $n-m \geq l \geq 1$, for whom $(\lambda_i - \chi^*)^2 \leq \lambda_i^2$ holds. Hence $\Omega^*(\sigma, \Lambda) = \chi^*\left(\sigma, \lambda_{(n-m+1)}\right)$.[12] This states more formally that, first, only the spatial swing player $(n-m+1)$ *may have* an influence on the outcome and, second, he actually *has* an influence only for particular preference constellations (here for $\lambda_{(n-m+1)} \in (0, \frac{1}{2}\sigma)$).

This calls for a further refinement of spatial swings. Namely, we say that player i has a *strict* (σ, Λ)-spatial swing in winning coalition S or is *strictly* (σ, Λ)-*pivotal* if his ideal policy outcome λ_i affects the agenda setter's optimal policy proposal $\chi^*(\sigma, \Lambda)$, i. e.$\partial\chi^*(\sigma, \Lambda)/\partial\lambda_i > 0$.[13] Clearly, a strict spatial swing implies a spatial swing. Note that at most one—and possibly no—voter can have a strict spatial swing for any given (σ, Λ)-realization. There can be lucky players who get more utility from the outcome than the swing player. This illustrates that being powerful does not per se imply particular success.

Considering a particular (σ, Λ)-combination, the players who are not (χ, Λ)-spatially pivotal for the agenda setter's optimal proposal $\chi = \chi^*(\Lambda, \sigma)$, do never influence the policy outcome for individually rational voting. They can be compared to excess players of a winning coalition in the coalitional form framework. A

[10] Identity of two or more players' ideal points has zero probability for a continuous distribution of Λ. This case will therefore be neglected in the following.

[11] One may assume small costs of being rejected for agenda setter A to ensure uniqueness of A's proposal in the last sub-case. There are, depending on Λ, multiple subgame perfect equilibria corresponding to the same unique equilibrium proposal by agenda setter A. We focus on (χ^*, Λ)-IR coalitions.

[12] Note that the ideal point $\lambda_{(n-m+1)}$ of the pivotal player is unique. In qualified majority voting there are two potential pivotal players but agenda setting makes the equilibrium unique.

[13] The possible event for which $\chi^*(\cdot)$'s derivative is not defined has zero probability and is therefore neglected.

player who has a spatial swing in $N_{\chi \succsim 0}$ but does not affect the agenda setter's proposal χ, i.e. has no *strict* spatial swing, is more like an inferior player in the coalitional form framework: He seems powerful as long as strategic considerations of decision-making are left out of the picture. Taking strategic interaction into account, he has no more power than true excess or dummy players.

As mentioned, very particular (σ, Λ)-combinations are of little interest for *a priori* power measurement. What matters is the a priori probability that a player ends up having power. This clearly depends on the distributional assumptions on $\tilde{\Lambda}$ and $\tilde{\sigma}$ that one makes. Forgetting for the moment the particular assumptions we have made above, it is generally useful to single out those players for which the necessary condition for influencing the outcome holds under almost no realization of ideal points, i. e. who almost never have a $(\chi^*(\sigma, \Lambda), \Lambda)$-spatial swing.

Definition 3 A player is called spatially dummy if

$$P\left\{\tilde{\lambda}_i = \tilde{\lambda}_{(n-m+1)}\right\} = 0.$$

$P\{\tilde{\lambda}_i = \tilde{\lambda}_{(n-m+1)}\} > 0$ is, however, not sufficient for a priori power in our agenda setting game. A player's spatial swing must, in addition, have positive probability of making a difference, i.e. of actually being a strict spatial swing. Player $(n - m + 1)$ has a strict spatial swing in the above setting if

$$0 < \lambda_{(n-m+1)} < \frac{1}{2}\sigma.$$

It is now in the spirit of the inferior player definition of Napel and Widgrén (2001) to define:

Definition 4 A player is called spatially inferior if

$$P\left\{\tilde{\lambda}_i = \tilde{\lambda}_{(n-m+1)} \wedge 0 < \tilde{\lambda}_{(n-m+1)} < \frac{1}{2}\tilde{\sigma}\right\} = 0$$

The probabilistic approach to the measurement of power in coalitional form games (cp. Straffin et al. (1977)) can straightforwardly be extended to measure a priori power in voting games with random spatial preferences. Namely, one measures a player's power as the probability of having a 'powerful' position. Building immediately on the more demanding notion of power embodied by strict spatial swings, this yields:

Definition 5 Consider a spatial voting game defined by $\left(\mathrm{N}, \mathrm{A}, \mathcal{W}, \mathrm{F}_{\tilde{\Lambda}}, \mathrm{F}_{\tilde{\sigma}}\right)$ and agenda setting as specified above. Then, the Strict Strategic Power Index (SSPI) ξ is defined by (i $\in$ N)

$$\xi_i \equiv P\left\{\tilde{\lambda}_i = \tilde{\lambda}_{(n-m+1)} \wedge 0 < \tilde{\lambda}_{(n-m+1)} < \frac{1}{2}\tilde{\sigma}\right\}.$$

Recall that in coalitional form voting games, players' preferences are not explicitly modelled. It is then a standard assumption to consider any ordering of players as equally probable and to attribute a swing to the $n-m+1$th player in a given ordering.[14] This produces the *Shapley-Shubik index (SSI)* ϕ. It may simply be expressed as

$$\phi_i \equiv P\left\{\tilde{\lambda}_i = \tilde{\lambda}_{(n-m+1)}\right\}$$

under the condition that the joint distribution of Λ makes all orderings equally probable.

The assumption of independent identically distributed (i. i. d.) uniform distributions satisfies this condition. Therefore, in the above setting, the SSPI can be expressed in terms of the Shapley-Shubik index ϕ_i in the subgame among voters:

$$\begin{aligned}\xi_i &= P\left\{\tilde{\lambda}_i = \tilde{\lambda}_{(n-m+1)}\right\} P\left\{0 < \tilde{\lambda}_{(n-m+1)} < \frac{1}{2}\tilde{\sigma} \middle| \tilde{\lambda}_i = \tilde{\lambda}_{(n-m+1)}\right\} \\ &= \phi_i \, P\left\{0 < \tilde{\lambda}_{(n-m+1)} < \frac{1}{2}\tilde{\sigma} \middle| \tilde{\lambda}_i = \tilde{\lambda}_{(n-m+1)}\right\}.\end{aligned}$$

In order to calculate the SSPI for a given spatial voting game with i.i.d. random preferences, the following result is useful:

Lemma Consider the i. i. d. random variables $\hat{\lambda}_1, \ldots, \hat{\lambda}_n$ with density $f_{\hat{\lambda}}$ and cumulative distribution function $F_{\hat{\lambda}}$. Let $\hat{\lambda}_{(p)}$ denote the pth order statistic of these n random variables, i. e. the (random) pth smallest value of $\hat{\lambda}_1, \ldots, \hat{\lambda}_n$. *Then*

$$\begin{aligned}F_{\hat{\lambda}_i = \hat{\lambda}_{(p)}}(x) &\equiv P(\hat{\lambda}_i \leq x \wedge \hat{\lambda}_i = \hat{\lambda}_{(p)}) \\ &= \int_0^x \binom{n-1}{p-1} F_{\hat{\lambda}}(s)^{p-1} \left[1 - F_{\hat{\lambda}}(s)\right]^{n-p} f_{\hat{\lambda}}(s)\, ds.\end{aligned}$$

Proof For both $\hat{\lambda}_i$ and $\hat{\lambda}_{(p)}$ to be equal to x, exactly $p-1$ random variables $\hat{\lambda}_j$, $j \neq i$, have to be no greater than x and the other $n-p$ random variables $\hat{\lambda}_j$, $j \neq i$, have to be no smaller than x (see e.g. Arnold et al. (1992)). There are $\binom{n-1}{p-1}$ permutations of $\hat{\lambda}_j$, $j \neq i$, that satisfy this requirement. Therefore[15]

---

[14] Equivalently, the mth player in a given order can be considered—this is just a matter of convention. A truly alternative assumption is to consider any coalition equally probable and any player in a given coalition as equally likely to leave. This leads to the *Banzhaf index.*

[15] If the $\hat{\lambda}_i$ are $U(0,1)$-distributed, this means that $\hat{\lambda}_{(p)}$ is Beta-distributed with parameters $(p, n-p+1)$.

$$
\begin{aligned}
&P(\hat{\lambda}_i \leq x \wedge \hat{\lambda}_i = \hat{\lambda}_{(p)}) \\
&\quad = \int_0^x \binom{n-1}{p-1} P(\lambda_1, \ldots, \lambda_{p-1} \leq s)\, P(\lambda_{p+1}, \ldots, \lambda_n \geq s) f_{\hat{\lambda}}(s)\, ds \\
&\quad = \int_0^x \binom{n-1}{p-1} F_{\hat{\lambda}}(s)^{p-1} \left[1 - F_{\hat{\lambda}(s)}\right]^{n-p} f_{\hat{\lambda}}(s)\, ds \\
&\quad = \frac{1}{n} \underbrace{\int_0^x n \binom{n-1}{p-1} F_{\hat{\lambda}}(s)^{p-1} \left[1 - F_{\hat{\lambda}(s)}\right]^{n-p} f_{\hat{\lambda}}(s)\, ds}_{P(\hat{\lambda}_{(p)} \leq x)} .
\end{aligned}
$$

□

Specifically, let us consider i.i.d. U(0,1) random variables $\hat{\lambda}_1, \ldots, \hat{\lambda}_n$, their pth order statistic $\hat{\lambda}_{(p)}$, and $\hat{\Pi}$ (independently $U(\frac{-\alpha}{(\beta-\alpha)}, \frac{\frac{1}{2}\beta-\alpha}{\beta-\alpha})$-distributed with density $f_{\hat{\Pi}}$). With this we get

$$
\begin{aligned}
&P(\hat{\lambda}_i \leq \hat{\Pi} \wedge \hat{\lambda}_i = \hat{\lambda}_{(p)}) \\
&\quad = \int_{-\infty}^{\infty} P(\hat{\lambda}_i \leq x \wedge \hat{\lambda}_i = \hat{\lambda}_{(p)}) f_{\hat{\Pi}}(x)\, dx \\
&\quad = \frac{1}{n} \int_{-\infty}^{\infty} P(\hat{\lambda}_{(p)} \leq x) f_{\hat{\Pi}}(x)\, dx \\
&\quad = \frac{1}{n} \int_{-\infty}^{\frac{-\alpha}{\beta-\alpha}} 0 \cdot f_{\hat{\Pi}}(x)\, dx + \frac{1}{n} \int_{\frac{-\alpha}{\beta-\alpha}}^{\frac{\frac{1}{2}\beta-\alpha}{\beta-\alpha}} P(\hat{\lambda}_{(p)} \leq x) f_{\hat{\Pi}}(x)\, dx + \frac{1}{n} \int_{\frac{\frac{1}{2}\beta-\alpha}{\beta-\alpha}}^{\infty} 1 \cdot 0\, dx \\
&\quad = \int_{\frac{-\alpha}{\beta-\alpha}}^{\frac{\frac{1}{2}\beta-\alpha}{\beta-\alpha}} \left[\int_0^x \binom{n-1}{p-1} s^{p-1} [1-s]^{n-p} ds \right] \cdot \frac{2(\beta-\alpha)}{\beta}\, dx
\end{aligned}
$$

With $p \equiv n - m + 1$ and the above distributional assumptions, we can now derive the explicit functional form of the SSPI in our example agenda setting model:

$$\xi_i = P\left\{\tilde{\lambda}_i = \tilde{\lambda}_{(p)} \wedge 0 < \tilde{\lambda}_{(p)} < \frac{1}{2}\tilde{\sigma}\right\}$$

$$= P\left\{\tilde{\lambda}_{(p)} < \frac{1}{2}\tilde{\sigma} \wedge \tilde{\lambda}_i = \tilde{\lambda}_{(n-m+1)}\right\} - P\left\{\tilde{\lambda}_{(p)} < 0 \wedge \tilde{\lambda}_i = \tilde{\lambda}_{(p)}\right\}$$

$$= \int_{-\infty}^{\infty} P\left\{\hat{\lambda}_{(p)} < x \wedge \hat{\lambda}_i = \hat{\lambda}_{(p)}\right\} f_{\hat{\Pi}}(x)\, dx$$

$$- P\left\{\hat{\lambda}_{(p)} < \frac{-\alpha}{\beta-\alpha} \wedge \hat{\lambda}_i = \hat{\lambda}_{(p)}\right\}$$

$$= \int_{\frac{-\alpha}{\beta-\alpha}}^{\frac{\frac{1}{2}\beta-\alpha}{\beta-\alpha}} \left[\int_0^x \binom{n-1}{p-1} s^{p-1}[1-s]^{n-p} ds\right] \cdot \frac{2(\beta-\alpha)}{\beta}\, dx$$

$$- \int_0^{\frac{-\alpha}{\beta-\alpha}} \binom{n-1}{p-1} s^{p-1}[1-s]^{n-p} ds.$$

Let us finally illustrate the SSPI, and also the difference between the inferior player axiom and the spatial inferiority condition, with an example:

Example 1 Consider the 3-person coalitional form game with $N = \{A, B, C\}; \mathcal{W} = \{\{A, B\}; \{A, C\}; \{A, B, C\}\}$. Given a uni-dimensional policy space, a natural model for this coalitional form game is uni-dimensional 3-person spatial agenda setting game. Suppose that $Q = 0$, $\alpha = -\frac{1}{3}$ and $\beta = 1$, implying that with probability $\frac{1}{4}$ there is no overlap between the agenda setter's and a voter's political interests. Note that this is exactly the same example as above but modified into a spatial setting. Players B and C are inferior. Player A is in the position to make take-it-or-leave-it offers to them, and hence the natural agenda setter. Let us denote the ideal point of A by a and the ideal points of B and C by b and c respectively. Suppose as above that $\tilde{a} \sim U(0, 1)$, $\tilde{b} \sim U\left(-\frac{1}{3}, 1\right)$ and $\tilde{c} \sim U\left(-\frac{1}{3}, 1\right)$. Re-scaling yields $\hat{a} \sim U\left(\frac{1}{4}, 1\right)$, $\hat{b} \sim U(0, 1)$, $\hat{c} \sim U(0, 1)$ and $\hat{\Pi} \sim U\left(\frac{1}{4}, \frac{5}{8}\right)$. Note that in this simple game we have $n = 2$ and $m = 1$. This implies

$$\begin{aligned}
\xi_B = \xi_C &= \int_{\frac{1}{4}}^{\frac{5}{8}} \left[\int_0^x \binom{1}{1} s\, ds \right] \cdot \frac{8}{3}\, dx - \int_0^{\frac{1}{4}} \binom{1}{1} s\, ds \\
&= \int_{\frac{1}{4}}^{\frac{5}{8}} \frac{8}{3} \cdot \frac{1}{2} x^2\, dx - \frac{1}{32} \\
&= \frac{4}{9} \left[\left(\frac{5}{8}\right)^3 - \left(\frac{1}{4}\right)^3 \right] - \frac{1}{32} \\
&= \frac{9}{128} \approx 0.0703.
\end{aligned}$$

This is less than half of the SSI, which gives $\frac{1}{6}$ for inferior players.

To further illustrate the difference between the SSPI and SSI let us first remove the range in which there are no gains from "exchange", i. e. set $\alpha = 0$. This implies $\hat{a} \sim U(0, 1)\hat{b} \sim U(0, 1)$, $\hat{c} \sim U(0, 1)$ and $\hat{\Pi} \sim U\left(0, \frac{1}{2}\right)$. Doing the same calculations as in the example above we get

$$\begin{aligned}
\xi_B = \xi_C &= \int_0^{\frac{1}{2}} \left[\int_0^x \binom{1}{1} s\, ds \right] \cdot \frac{2}{1}\, dx \\
&= \frac{1}{3} \left(\frac{1}{2}\right)^3 = \frac{1}{24} \approx 0.0417.
\end{aligned}$$

When it becomes more likely that a proposal is accepted, it also becomes more likely that the ideal point of the agenda setter A is accepted. Inefficiency, i. e. $\alpha < 0$, benefits the voters since it complicates strategic agenda setting. This is not the case when the agenda setter does not act strategically. Then the extent of status quo bias has no role. To see this let us assume that the agenda setter becomes like one of the voters and is acting non-strategically by always proposing $\chi = \lambda_{(n-m+1)}$.[16] In the example above, this means that the assumed agenda setter A is able to pass her ideal point in four ideal point permutations: $(\lambda_B, \lambda_A, \lambda_C), (\lambda_C, \lambda_A, \lambda_B)$, $(\lambda_B, \lambda_C, \lambda_A)$, $(\lambda_C, \lambda_B, \lambda_A)$. Note that this is independent of the value of α since in this case IR-constraints do not affect agenda setting. Players B and C are able to make a change in $(\lambda_A, \lambda_B, \lambda_C)$ and $(\lambda_A, \lambda_C, \lambda_B)$, respectively. Hence we get $\left(\frac{2}{3}, \frac{1}{6}, \frac{1}{6}\right)$, the SSI. The SSPI—by adding strategic agenda setting to a spatial voting model—yields

[16] Alternatively we can think that the agenda setter is really like a voter and a proposal is made by an intelligent benevolent machine after the players have told it their ideal points.

something reminiscent to the SSI with the degree of similarity determined by various factors. The model demonstrates that inefficiencies in decision making, as measured by α, have significant impact on power if it is understood as the ability to make a difference.[17] Hence we get different power distributions when we let the value of α vary.[18] This is in a sense trivial: If one analyses how spatial preferences affect power, the domain and distribution of preferences matter.

3 Concluding remarks

In spatial voting games the individual rationality constraints above determine what kind of proposals will be accepted. Players' rates of acceptance are thus determined by the relative locations of voters' and the agenda setter's ideal points. Moreover, the agenda setter is assumed to act strategically. Strategic aspects of coalition formation were introduced into coalitional form games in Napel and Widgrén (2001) by distinguishing between inferior and non-inferior players. The implications on players' power were discussed more in depth in Napel and Widgrén (2002). In this chapter, following this tradition, we have constructed a strategic power index, which has spatial preferences and strategic agenda setting as its main building blocks. Earlier work in this field is still preliminary. In Steunenberg et al. (1999), a different strategic power index is introduced. This measure, contrary to what we propose here, defines power as proximity between one's ideal point and the outcome of the game. But, proximity may be due to luck and, indeed, in this chapter we demonstrate that under strategic agenda setting players whose ideal points are located close to the outcome tend to have luck, not power. The pivotal player is the player who exerts power, although a winner's curse often arises in terms of proximity. In fact, when the pivotal player has an effect to the outcome she gains the least among the players in a winning coalition.

In this chapter, we have proposed a new strategic power index for spatial voting games. Our model has several restrictions like uni-dimensionality and a specific sequential game form. We feel, however, that we have opened an avenue for a new type of power measurement literature and further research should follow.

---

[17] The values of the SSI and the SSPI are comparable as probabilities. The values of the SSPI shed some light how much difference strategic agenda setting makes to the SSI under different assumptions of the domains of preference distributions. Note, however, that the purpose of this chapter is not a beauty contest between the SSI and the SSPI. Our attempt is to assess the relationship between spatial preferences and power. As a special case we get the SSI.

[18] Strictly speaking we let the ratio $\frac{\alpha}{\beta}$ vary. This ratio affects the re-scaling presented above.

References

Arnold, B. C., Balakrishnan, N., & Nagaraja, H. N. (1992). *A first course in order statistics, Wiley series in probability and mathematical statistics*. New York: Wiley.

Banzhaf, J. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review*, *19*, 317–343.

Bolger, E. (1980). A class of power indices for voting games. *International Journal of Game Theory*, *9*, 217–232.

Deegan, J., & Packel, E. W. (1979). A new index of power for simple n-person games. *International Journal of Game Theory*, *7*, 113–123.

Felsenthal, D., & Machover, M. (1998). *The measurement of voting power—Theory and practice, problems and paradoxes*. Cheltenham: Edward Elgar.

Holler, M.J., & Owen, G. (Eds.). (2001). *Power indices and coalition formation*, Dordrecht: Kluwer.

Holler, M., & Packel, E. (1983). Power, luck and the right index. *Zeitschrift für Nationalökonomie*, *43*, 21–29.

Napel, S., & Widgrén, M. (2001). Inferior players in simple games. *International Journal of Game Theory*, *30*, 209–220.

Napel, S., & Widgrén, M. (2002). The power of an inferior player. In M. J. Holler, H. Kliemt, D. Schmidtchen& M. E. Streit (Eds.), *Jahrbuch für Neue Politische Ökonomie*. Tübingen: Mohr-Siebeck, *20*, 259–273.

Napel, S., & Widgrén, M. (2004). Power measurement as sensitivity analysis—a unified approach. *Journal of Theoretical Politics*, *16*(4), 517–538.

Shapley, L. S., & Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review*, *48*, 787–792.

Steunenberg, B., Schmidtchen, D., & Koboldt, C. (1999). Strategic power in the European union: Evaluating the distribution of power in policy games. *Journal of Theoretical Politics*, *11*, 339–366.

Straffin, P. D., Jr. (1977). Homogeneity, independence and power indices. *Public Choice*, *30*, 107–118.

On the Nucleolus as a Power Index

Maria Montero

1 Introduction

Consider the classical problem of dividing a dollar by majority rule. There are n members in the voting body and a voting rule characterized by a set of winning coalitions. How powerful is each member of this voting body? A measure of power is the expected share of the budget for each player. This concept of power is what Felsenthal and Machover (1998) call P-power. They point out that the outcome of the bargaining process will not generally be deterministic,[1] and the index of P-power will be an average of the possible outcomes, weighted by their probability.

In this context, the most widely accepted measure of power is the Shapley value, introduced by Shapley in 1953 as a measure of the expected payoff from playing a general cooperative game. The Shapley value can be seen as a weighted average of several possible outcomes (the simplest possibility is Shapley's original story of players entering randomly into a room and receiving their marginal contribution to the value of the existing group). Dubey (1975), Roth (1977a, b) and Laruelle and Valenciano (2007) give further axiomatic support to the Shapley value as a measure of power in divide-the-dollar games.

In contrast, the nucleolus (introduced by Schmeidler in 1969, see also Maschler (1992) answers a different type of question: what is the most stable way of dividing

This chapter has benefitted from discussion with Manfred Holler and the comments of two referees. An earlier version of this chapter has been published in Homo Oeconomicus 22(4)

[1] Consider the simple majority game with three players. If there was a unique deterministic outcome, symmetry points to the grand coalition with every player receiving $\frac{1}{3}$. However, this outcome seems too fragile. If we accept that a two-player coalition will eventually form, symmetry dictates that each of the three possible coalitions will be equally likely.

M. Montero (✉)
School of Economics, University of Nottingham, University Park,
Nottingham NG7 2RD, UK
e-mail: maria.montero@nottingham.ac.uk

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_15,

the dollar between the n players? Unlike the Shapley value, the nucleolus seems to presuppose the formation of the grand coalition and therefore seems inappropriate as a measure of P-power. The nucleolus is the most stable outcome given that the grand coalition forms, not the average of several possible outcomes which may not involve the grand coalition at all. Accordingly the nucleolus is not even mentioned by Felsenthal and Machover (1998), though it appears in Pajala's (2002) literature list.

Nevertheless, this chapter will discuss an alternative interpretation of the nucleolus and argue that it can compete with the Shapley value as a measure of P-power. The most solid reason lies in the field of noncooperative foundations, but the chapter will also discuss some general properties of the Shapley value and the nucleolus.

The nucleolus will find support in the Baron and Ferejohn (1989) model. This is not a model introduced in order to provide noncooperative foundations to any particular solution concept, and it is fairly popular with political scientists.[2] The Shapley value finds little or no support in this model. The chapter will also discuss why the existing literature on noncooperative foundations of the Shapley value is either not applicable or not fully convincing for majority games.

The question arises of whether the nucleolus is a good power index if we abstract from the bargaining process. A possible (and solid) reason why the nucleolus has been ignored as a power index is that it may assign zero to players who are not dummies; this seems counterintuitive and indeed Felsenthal and Machover (1998) include "vanishing only for dummies" as one of the postulates any power index must obey. Another postulate that Felsenthal and Machover consider essential is that a power index respects dominance, i.e., exactly mirrors the ranking of players by desirability. The nucleolus respects dominance only weakly: it is possible for two players to get the same payoff even though one is more desirable than the other. On the other hand, the requirements of vanishing only for dummies and respecting dominance are incompatible with core selection. Moreover, these properties of the nucleolus can be justified if we interpret the nucleolus as a system of competitive prices. Finally, the nucleolus does better than the Shapley value with respect to the added blocker postulate.

The remainder of the chapter is organized as follows. Section 2 contains some preliminaries on majority games and the nucleolus. Section 3 provides an alternative interpretation of the nucleolus as a system of competitive prices for the players. Section 4 illustrates how those competitive prices can arise as the equilibrium of a natural modification of the Baron-Ferejohn model. Section 5 discusses some properties of the nucleolus as a power index, and Sect. 6 concludes.

[2] A search in Thomson Reuters's "Web of Knowledge" conducted in January 2012 reveals that the paper has been cited 389 times; about 40 % of those citations appeared in political science journals.

2 Preliminaries

2.1 Majority Games

Let $N = \{1, \ldots, n\}$ be the set of players. $S \subseteq N (S \neq \emptyset)$ represents a generic coalition of players, and $v : 2^n \to \mathbb{R}$ with $v(\emptyset) = 0$ is a function that assigns to each coalition the total payoff its members can divide among themselves. The function v is called *characteristic function*. In the context of *majority* (also called *simple*) games, $v(S) \in \{0, 1\}$ for all $S \subseteq N, v(\emptyset) = 0, v(N) = 1,$ and $v(T) = 1$ whenever $v(S) = 1$ for some $S \subseteq T \subseteq N$. A coalition S is called *winning* iff $v(S) = 1$ and *losing* iff $v(S) = 0$. It is called *minimal winning* iff $v(S) = 1$ and $v(T) = 0$ for all T such that $T \subset S$. The set of winning coalitions is denoted by W; this set contains the same information as the function v. The set of minimal winning coalitions is denoted by W^m. A player i such that $v(S \cup i) = v(S)$ for all S is called a *dummy player*. A player who belongs to all winning coalitions is called a *veto player* or a *blocker*.

A simple game is *proper* iff $v(S) = 1$ implies $v(T) = 0$ for all $T \subset N \backslash S$. It is *constant-sum* iff $v(S) + v(N \backslash S) = 1$ for all $S \subseteq N$. It is a *weighted majority game* iff there exist n nonnegative numbers (weights) $w_1, \ldots, w_n$ and a positive number q such that $v(S) = 1$ if and only if $\sum_{i \in S} w_i := w(S) \geq q$. We will denote a weighted majority game by $[q; w_1, \ldots, w_n]$. The pair $[q; w]$ is called a *representation* of the game v. A weighted majority game has many possible representations, but not all of them are equally convenient. A representation w is called *normalized* iff $w(N) = 1$; it is *homogeneous* iff $w(S) = q$ for all $S \in W^m$. Not all weighted majority games admit a homogeneous representation. A weighted majority game admitting a homogeneous representation is called a *homogeneous game*.[3]

2.2 The Nucleolus

Let (N, v) be a majority game and $x = (x_1, \ldots, x_n)$ be an *imputation*, that is, a payoff vector with $x_i \geq v(i)$ and $x(N) = v(N)$. For any coalition S the value $e(S, x) = v(S) - x(S)$ is called the *excess* of S at x.

For any imputation x let $S_1, \ldots, S_{2^{|N|}-1}$ be an ordering of the coalitions for which $e(S_l, x) \geq e(S_{l+1}, x)$ for all $l = 1, \ldots, 2^{|N|} - 1$ and let $E(x)$ be the vector of excesses defined by $E_l(x) = e(S_l, x)$ for all $l = 1, \ldots, 2^{|N|} - 1$. We say that $E(x)$ is *lexicographically less* than $E(y)$ if $E_l(x) < E_l(y)$ for the smallest l for which $E_l(x) \neq E_l(y)$. The *nucleolus* is the set of imputations x for which the vector $E(x)$

[3] Homogeneous representations are preferable because they give a more accurate description of the situation. For example, games $[5; 4, 3, 2]$ and $[2; 1, 1, 1]$ are identical in terms of the characteristic function, but only the second representation is homogeneous. This representation reflects the fact that all players are in a symmetric position.

is lexicographically minimal. Schmeidler (1969) shows that the nucleolus consists of a unique imputation. It is contained in the classical bargaining set (Davis and Maschler 1967) and in the kernel (Davis and Maschler 1965).

The excess of coalition S at x is a measure of how dissatisfied coalition S is with imputation x. We can think of the excess as a measure of how likely S would be to depart from the grand coalition. The nucleolus minimizes the maximum excess, and thus is one of the (possibly many) solutions[4] of the following linear programming problem[5]

$$\begin{aligned} &\min e \\ &\text{s.t. } x(S) + e \geq 1 \text{for all } S \in W \\ &x(N) = 1 \\ &x_i \geq 0 \text{ for all } i \in N. \end{aligned}$$

An important property of the nucleolus is related to the fact that it is a solution to this linear programming problem. To present this property, we need some definitions.

For every majority game v and every payoff vector x, let $b_1(x, v)$ be the set of those $S \subseteq N$ for which $\max\{v(S) - x(S) : S \subseteq N\}$ is attained and $b_0(x) = \{\{i\} : x_i = 0\}$.

Let $\mathcal{C}$ be a collection of nonempty subsets of N. We say that the collection is *balanced* iff there exist strictly positive numbers $(\lambda_S)_{S\in\mathcal{C}}$ such that, for each $i \in N, \sum_{S\ni i} \lambda_S = 1$. The numbers $(\lambda_S)_{S\in\mathcal{C}}$ receive the name of *balancing weights*.

Proposition 1 (*Kohlberg* 1971) *Let v be a majority game. If x is the nucleolus of v, then there is a subset $b'_0(x)$ of $b_0(x)$ such that $b'_0(x) \cup b_1(x, v)$ is balanced.*

This property will play an important role in the next two sections.

3 The Nucleolus as a Competitive System of Prices

The important property of the balancing weights in the previous section is not that they add up to 1 for each player, but that they add up to the same constant for each player who gets a positive payoff. We can change this constant by rescaling the weights to obtain another set of weights λ'_S. In particular, suppose we rescale the weights in such a way that $\sum_{S\in b_1(x,v)} \lambda'_S = 1$. The weights $(\lambda'_S)_{S\in b_1(x,v)}$ can be interpreted as the probabilities of coalition S forming (cf. (Albers 1974, p. 5)). Then each player with $x_i > 0$ will be in the final coalition with the same

[4] Derks and Kuipers' have developed a program to compute the nucleolus, which can be found at Jean Derks' homepage (http://www.personeel.unimaas.nl/Jean-Derks/). The algorithm is explained in Derks and Kuipers (1997).

[5] This will be the case even if there is a player i with $v(i) = 1$.

probability (which turns out to be precisely the total payoff x assigns to a coalition of maximum excess), and a player with $x_i = 0$ appears in the final coalition no more often than one with $x_i > 0$. This interpretation of the balancing weights doesn't seem widespread - indeed Albers dismisses it in the related context of balanced aspirations.

Why interpret the weights $(\lambda'_S)_{S \in b_1(x,v)}$ as the probabilities of each coalition forming? Suppose the imputation $x := (x_1, \ldots, x_n)$ is related to a system of prices players charge for their participation in a coalition. Which sets of prices are stable? We can make two assumptions with respect to what happens when a coalition forms:

1. A privileged player i (the proposer) selects a coalition S, pays each player $j \in S \setminus \{i\}$ the price x_j and pockets the residual, which will generally be higher than x_i. In this case player i will choose S to solve the following problem

$$\max_{\substack{S \in W \\ S \ni i}} \; 1 - \sum_{j \in S \setminus \{i\}} x_j$$

 This problem is equivalent to $\max_{S \in W, S \ni i} 1 - \sum_{j \in S} x_j$. In other words, given a price vector x player i always proposes a coalition of maximum excess containing him. Because he only pays the others x_j and keeps the whole excess, he wants to maximize that excess. The set $b_1(x, v)$ becomes prominent (though in general not all players will belong to one of the coalitions in $b_1(x, v)$, if x is the nucleolus all players do). If x is the nucleolus it is reasonable to assume that only coalitions in $b_1(x, v)$ will form.
2. Alternatively, we can assume that no player is privileged and that if a coalition S forms, the players in S will divide the payoff proportionally to x. Again, coalitions in $b_1(x, v)$ emerge as the most profitable and it is reasonable to assume that they will form. This is because both the surplus above $\sum_{j \in S} x_j$ and the share of the surplus a player receives are maximized for coalitions of maximum excess. If x is the nucleolus, *ex post* payoff division will correspond to a balanced aspiration (see Cross 1967).

In any of the two cases we will have identified the coalitions that may form given the system of prices. Each of these coalitions will form with a certain probability. If, *for all possible probability distributions* over the set of coalitions, there is a group of players that belongs to the final coalition more often than another group of players, the price system is not stable. There is an "excess demand" for some players and their price should rise at the expense of some players who belong to the final coalition less often. The only exception is the case in which the players who are less often in the final coalition are already getting 0. On the other hand, if we can assign probabilities to the coalitions so that all players who get a positive payoff are in the final coalition with equal probability, we have a competitive price system.

Let the system of prices coincide with the nucleolus. As we have seen, in any of the two cases above the coalitions that will form given these prices are the ones with maximum excess. In either case, since the set of coalitions of maximum excess is balanced, we can assign probabilities to them such that all players with positive payoff are in the coalition with the same probability, and players who are getting 0 are no more often than other players in the final coalition.

Example 1 $[5;2,2,2,1,1,1]$. The nucleolus is $\left(\frac{2}{9},\frac{2}{9},\frac{2}{9},\frac{1}{9},\frac{1}{9},\frac{1}{9}\right)$. Given this, a large player will be indifferent between proposing a coalition of type [221] and [2111]. A small player will also be indifferent between them. There are nine coalitions of the first type and three of the second type (twelve in total). A large player is in seven coalitions, and a small player is in six. Thus, if all coalitions were equally likely the large players would be in excess demand. However, we can assign probabilities to coalitions such that all players are in the final coalition with probability $\frac{5}{9}$ (the total nucleolus payoff of a coalition of maximum excess). If we assign $\frac{2}{27}$ to type [221] and $\frac{1}{9}$ to [2111] this will be the case.

The Shapley value of this game is $\left(\frac{7}{30},\frac{7}{30},\frac{7}{30},\frac{3}{30},\frac{3}{30},\frac{3}{30}\right)$. Given these prices, all players will propose a coalition of type [2111]. A large player is in only one of those, whereas a small player is in all three. Thus, regardless of what probabilities we assign to the coalitions each small player will be in the final coalition with probability 1; each large player will be (on average) with probability $\frac{1}{3}$. There is a sense in which the small players are more in demand and should raise their price.

Example 2 $[5;3,2,2,1]$. The nucleolus is $\left(\frac{1}{3},\frac{1}{3},\frac{1}{3},0\right)$. Coalitions [32], [221] and others in which we add player [1] to a minimal winning coalition would be optimal. We can find weights ($\frac{1}{3}$ for each minimal winning coalition) such that all players who get a positive payoff are in the final coalition with the same probability ($\frac{2}{3}$), and the player who gets a zero payoff is in the final coalition with a smaller probability ($\frac{1}{3}$). Even though player 4 is “providing something for nothing” when he enters a coalition, he is not in excess demand.

The Shapley value is $\left(\frac{5}{12},\frac{3}{12},\frac{3}{12},\frac{1}{12}\right)$. Given these prices the cheapest coalition is [221], which does not include player [3]. We would expect player [3] to lower his price and players [2] to raise their price, as they are in demand by player [3].

4 Noncooperative Foundations

4.1 The Baron-Ferejohn model

Baron and Ferejohn (1989) influential paper introduced a legislative bargaining game based on Binmore (1987) modification of Rubinstein (1982) two-player bargaining game. In their chapter n symmetric players must divide a budget by simple majority. Each player has an equal chance of being recognized to be the proposer; once a proposer is recognized he proposes a division of the budget. The rest of players then vote “yes” or “no”; if a majority of the players supports the proposal then it is implemented and the game ends; otherwise we come back to the previous situation in which nature chooses a proposer, each player being chosen with equal probability. Baron and Ferejohn focus on stationary subgame perfect equilibria. In a stationary equilibrium, strategies do not depend on any elements of the history of the game other than the current proposal, if any. It is important to emphasize that Baron and Ferejohn’s model appeared in a political science journal; nothing seems to connect their paper with the field of noncooperative foundations.

In extending the model to general voting games we must choose whether to keep the recognition probabilities identical for all players, or to have asymmetric probabilities. If the game is a weighted majority game, we may want to select each player with a probability proportional to his number of votes (this is done by Baron and Ferejohn in one of their examples). This extension has a straightforward interpretation if players are parties, different number of votes correspond to different number of representatives, and each representative is selected to be the proposer with equal probability.

In Montero (2001), I extend the Baron-Ferejohn model to any proper simple game, and show that the nucleolus can always be obtained as the unique equilibrium expected shares in the Baron-Ferejohn game, provided that the recognition probabilities coincide with the nucleolus. Since the recognition probabilities are itself a measure of bargaining power (an input of the game, which in principle need not be related to the voting rule), the nucleolus is a sort of self-confirming power index in this noncooperative game. As for other recognition probabilities, the nucleolus seems more likely to emerge as an equilibrium than the Shapley value.

Example 3 Consider the game $[5; 3, 2, 2, 1, 1]$. The nucleolus of this game is $\left(\frac{3}{9}, \frac{2}{9}, \frac{2}{9}, \frac{1}{9}, \frac{1}{9}\right)$. In the Baron-Ferejohn bargaining procedure with recognition probabilities $\theta = \left(\frac{3}{9}, \frac{2}{9}, \frac{2}{9}, \frac{1}{9}, \frac{1}{9}\right)$, the only stationary equilibrium expected payoff is precisely $\left(\frac{3}{9}, \frac{2}{9}, \frac{2}{9}, \frac{1}{9}, \frac{1}{9}\right)$.

The idea of the proof of this result is as follows. Expected equilibrium payoffs act as reservation prices: if a proposal is rejected, nature starts the game all over again and, since strategies are stationary, each player receives his equilibrium payoff. It is then a best response for a player to accept any offer that gives him at least his equilibrium payoff.

Given this vector of prices, it is optimal for the proposer to propose a coalition of maximum excess. In this example all minimal winning coalitions are of maximum excess: $\{1,2\}, \{1,3\}, \{1,4,5\}, \{2,3,4\}$ and $\{2,3,5\}$. This collection is balanced; a set of balancing weights is $\lambda_{\{1,2\}} = \lambda_{\{1,3\}} = \frac{1}{5}, \lambda_{\{1,4,5\}} = \frac{3}{5}, \lambda_{\{2,3,4\}} = \lambda_{\{2,3,5\}} = \frac{2}{5}$. Consider the following strategy for the proposer: if he belongs to S, he proposes S with probability λ_S, and offers each other player in S their price. Because $\sum_{S \ni i} \lambda_S = 1$, the proposer’s strategy is completely determined. Moreover, if all players follow these strategies, expected payoffs indeed coincide with $\left(\frac{1}{3}, \frac{2}{9}, \frac{2}{9}, \frac{1}{9}, \frac{1}{9}\right)$.

Consider for example player 1. With probability $\frac{1}{3}$ he is selected to be the proposer; he then proposes $\{1,2\}$ or $\{1,3\}$—offering the other player his price of $\frac{2}{9}$ and thus obtaining $1-\frac{2}{9}=\frac{7}{9}$—or alternatively $\{1,4,5\}$—offering each of the other two players $\frac{1}{9}$ and thus also obtaining $\frac{7}{9}$. With probability $\frac{2}{9}$ player 2 is selected to be the proposer; player 2 proposes $\{1,2\}$ with probability $\lambda_{\{1,2\}}=\frac{1}{5}$, and pays player 1 his price, $\frac{1}{3}$; the same applies to player 3. Each of players 4 and 5 is selected with probability $\frac{1}{9}$, proposes $\{1,4,5\}$ with probability $\lambda_{\{1,4,5\}}=\frac{3}{5}$ and offers $\frac{1}{3}$ to player 1. Player 1's expected payoff is then

$$\frac{1}{3}\left[1-\frac{2}{9}\right]+\left[\frac{4}{9}\frac{1}{5}+\frac{2}{9}\frac{3}{5}\right]\frac{1}{3}=\frac{1}{3}.$$

These strategies have the property that the probabilities of each coalition forming are proportional to the balancing weights, vindicating the interpretation of balancing weights as related to the probability of each coalition forming. Notice also that each player is in the final coalition with the same probability, in this case $\frac{5}{9}$. Thus Montero (2001) contains a justification of the arguments in the previous section in a strategic model of coalition formation in which players are free to propose any coalition with any payoff division.

If we consider arbitrary recognition probabilities, Kalandrakis (2006) has shown that any payoff vector can be obtained in equilibrium for *some* choice of the recognition probabilities. This result does not imply that all payoff vectors have equal merit as we will see in the examples below: some payoffs can be obtained for a broad range of probabilities whereas others can only be obtained for a single probability vector; some equilibria may be fragile and would not survive perturbations of the game.

The nucleolus seems to be more likely to arise as an equilibrium than the Shapley value. This is because, as a price vector, the nucleolus makes the proposer indifferent between several attractive coalitions, whereas the Shapley value usually induces strict preferences over coalitions. The following example illustrates this point.

Example 4 Consider the game $[3;2,1,1,1]$ and a protocol that selects player 1 with probability θ_1 and each other player with probability $\frac{1-\theta_1}{3}$. The nucleolus of this game is $\left(\frac{2}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}\right)$ and the Shapley value is $\left(\frac{1}{2}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right)$. The nucleolus can be obtained for any $\theta_1 \leq \frac{1}{2}$; the Shapley value is only obtained for $\theta_1=\frac{3}{5}$.

Given the price vector $\left(\frac{2}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}\right)$, each of the players with 1 vote is indifferent between proposing to player 1 and proposing to the other two players: in both cases the proposer pays a total of $\frac{2}{5}$. We can construct an equilibrium in which player 1 proposes to each other player with probability $\frac{1}{3}$, and each other player proposes to player 1 with probability λ, where λ can be found from player 1's expected payoff equation $\frac{2}{5} = \theta_1\left[1 - \frac{1}{5}\right] + (1 - \theta_1)\lambda\frac{2}{5}$. The solution to this equation, $\lambda = \frac{1-2\theta_1}{1-\theta_1}$, is between 0 and 1 for $\theta_1 \leq \frac{1}{2}$.

In contrast, given the price vector $\left(\frac{1}{2}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right)$ player 1 is overpriced and receives no proposals, and the expected payoff equation becomes $\frac{1}{2} = \theta_1\left[1 - \frac{1}{6}\right]$, which has only one solution.

Even if the Shapley value makes players indifferent between the relevant coalitions, it may be all but impossible to obtain as the following example of a game with a nonempty core illustrates (see Winter 1996; Banks and Duggan 2000).

Example 5 Consider the game $[3; 2, 1, 1]$ and a protocol that selects player 1 with probability θ_1 and each other player with probability $\frac{1-\theta_1}{2}$. The nucleolus of this game is $(1, 0, 0)$ and the Shapley value is $\left(\frac{2}{3}, \frac{1}{6}, \frac{1}{6}\right)$. The nucleolus is the only equilibrium payoff for $\theta_1 > 0$. The Shapley value may arise for $\theta_1 = 0$, but this equilibrium is not robust.

Let y_1 be the expected equilibrium payoff for player 1 and y_2 the expected equilibrium payoff for 2 and 3. Then expected payoff for player 1 is given by $y_1 = \theta_1[1 - y_2] + (1 - \theta_1)y_1$. For $\theta_1 > 0$, the solution of this equation together with $y_1 = 1 - 2y_2$ is $y_1 = 1$. For $\theta_1 = 0$ we have $y_1 = y_1$, so in principle any payoff including the Shapley value could be an equilibrium. However, if we introduce a discount factor δ arbitrarily close to 1 in order to ensure uniqueness of equilibrium,[6] the equation becomes $y_1 = \delta y_1$, and its only solution is $y_1 = 0$. Under this assumption *no* protocol implements the Shapley value. The same applies to all power indices based on marginal contributions or that give positive values to any player who is at least in one minimal winning coalition, like the Johnston (1978) and Deegan and Packel (1978) indices. Nohn (2010a, b) has recently emphasized this point and shown that it still holds in a more general model that allows recognition probabilities to depend on some elements of the history of play.

The Baron-Ferejohn model has been criticized because of the disproportionate advantage it gives to the proposer (see e.g. Harrington (1990)). However, it can be easily modified to eliminate this advantage, as Montero (2008) shows.

Even if the core is empty, the nucleolus may give a payoff of 0 to players that are not dummies. For example, in the game $[5; 3, 2, 2, 1]$ the nucleolus is $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0\right)$. This vector can be obtained as the unique expected equilibrium payoff in the Baron-Ferejohn model, but only if player 4 is never selected to be proposer. However, no power index is generally supported by natural protocols like the egalitarian or the proportional protocol.[7]

---

[6] Rubinstein (1982) introduced a discount factor in order to achieve uniqueness of subgame perfect equilibrium in the two-player bargaining game. In the general Baron-Ferejohn model with arbitrary majority games and recognition probabilities, Eraslan and McLennan (2006) show that all stationary subgame perfect equilibria must have the same payoffs for $\delta < 1$.

[7] This includes the modified nucleolus of Sudhölter (1996), which is a representation of all weighted majority games.

4.2 Discussion of Other Bargaining Models

An alternative noncooperative foundation for the nucleolus in majority games can be found in Young (1978). He shows that the nucleolus can be obtained as an equilibrium payoff in an asymmetric lobbying game where two lobbyists compete in order to buy the players' votes, and one of the lobbyists has substantially more resources than the other (see also Le Breton and Zaporozhets 2010). Unfortunately, the game becomes very difficult to solve if the two lobbyists move simultaneously and have equal resources.[8]

The Shapley value has some noncooperative foundations of its own. The most natural model, that of Gul (1989), is not applicable to simple games because it requires that any two players benefit (in terms of the Shapley value) from forming a bloc. Winter (1994) demand commitment model only applies to convex games and thus it is not applicable to simple games. Other noncooperative foundations of the Shapley value are formally applicable to simple games, but they give a special role to the grand coalition, which seems contradictory with the idea of a majority game: Hart and Mas-Colell (1996), Pérez-Castrillo and Wettstein (2001) and Laruelle and Valenciano (2008) require *all* players to agree with a proposal. Vidal-Puga (2008) allows only one coalition (not necessarily the grand coalition), and players must choose between joining it or become singletons; they are not allowed to wait, even though they may actually prefer to do so. Instead in the Baron-Ferejohn model players always prefer to be proposers rather than wait.

It seems paradoxical that, while the Shapley value is usually interpreted as an expected payoff of playing the game which unlike the nucleolus does not presuppose the grand coalition, the opposite happens in the corresponding implementations: Hart and Mas-Colell require the grand coalition to form in order to obtain the Shapley value while the nucleolus can be obtained in the Baron-Ferejohn model as a "value" without giving the grand coalition a prominent role (indeed the grand coalition is never formed in the absence of veto players).

5 Some Properties of the Shapley Value and Nucleolus

The nucleolus satisfies the following property: suppose, as in our previous discussion, that only coalitions of maximum excess given a price vector x form, and payoff division inside a coalition is proportional to this price vector. Let $\mathbf{S}_i$ be the set of coalitions of maximum excess to which i belongs. If $\mathbf{S}_i \subseteq \mathbf{S}_j$ and $\mathbf{S}_j \not\subseteq \mathbf{S}_i$, one can say that i depends on j, but j does not depend on i. In this case one would expect i to reduce his payoff in favor of j, unless $x_i = 0$. This property is called the *partnership condition* by Bennett (1983), and a very similar condition is postulated

[8] Young (1978) provides the result of calculations for two weighted majority games; these results do not coincide with any of the known power indices.

by Napel and Widgrén 2001 as a desirable property of a power index. The fact that the nucleolus has this property is clear from Kohlberg's result, as the following claim shows.

Claim 1 Let (N, v) be a simple game, x the nucleolus of v, and $\mathbf{S}_i = \{S \in b_1(x, v) : i \in S\}$. Then for any two players i and j, $\mathbf{S}_i \subseteq \mathbf{S}_j$ and $\mathbf{S}_j \nsubseteq \mathbf{S}_i$ implies $x_i = 0$.

Proof Suppose $\mathbf{S}_i \subseteq \mathbf{S}_j$ and $\mathbf{S}_j \nsubseteq \mathbf{S}_i$, but $x_i > 0$. Because x is the nucleolus, the set $b'_0(x) \cup b_1(x, v)$ must be balanced. Let $(\lambda_S)_{S \in b'_0(x) \cup b_1(x,v)}$ be a set of balancing weights. Because $x_i > 0$, $\sum_{S \in b_1(x,v), S \ni i} \lambda_S = 1$. But $\sum_{S \in b_1(x,v), S \ni i} \lambda_S < \sum_{S_1(x,v), S \ni j} \lambda_S = 1$, a contradiction □

The Shapley value does not seem to have an analogous property: if we consider the majority game with a veto player $[3; 2, 1, 1]$, players 2 and 3 clearly depend on 1 but still have a positive Shapley value.

A property enjoyed by the Shapley value but not by the nucleolus is the *symmetric gain/loss* property (see Laruelle and Valenciano (2001)). This property states that, if we compare a simple game v with the game v^*_S that results after deleting a minimal winning coalition $S \neq N$ from v, then the change in the Shapley value is the same for all players in S and for all players in $N \backslash S$. The following example illustrates this property:

Example 6 Consider the game $(5; 3, 2, 2, 1, 1)$. This game has the following minimal winning coalitions: $\{1, 2\}, \{1, 3\}, \{2, 3, 4\}, \{2, 3, 5\}, \{1, 4, 5\}$. Players 1, 2 and 3 belong to three of those, whereas players 4 and 5 belong only to two of those.

Now consider the modified game that has the same characteristic function except that $v(1, 4, 5) = 0$. The Shapley value of the original game is $\left(\frac{24}{60}, \frac{14}{60}, \frac{14}{60}, \frac{4}{60}, \frac{4}{60}\right)$; after deleting coalition $\{1, 4, 5\}$ from the set of winning coalitions the Shapley value changes to $\left(\frac{22}{60}, \frac{17}{60}, \frac{17}{60}, \frac{2}{60}, \frac{2}{60}\right)$. Thus, each of players 1, 4 and 5 have lost $\frac{2}{60}$. However, one may argue that coalition $\{1, 4, 5\}$ was crucial for players 4 and 5 but not for player 1. After the deletion of $\{1, 4, 5\}$, player 1 can form a coalition with either 2 or 3, whereas players 4 and 5 are now dependent on players 2 and 3. Indeed the nucleolus changes from $\left(\frac{3}{9}, \frac{2}{9}, \frac{2}{9}, \frac{1}{9}, \frac{1}{9}\right)$ to the (somewhat extreme but consistent with the partnership condition) $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0\right)$. Players 4 and 5 are seriously affected, but not so player 1.

Because the nucleolus is symmetric and always belongs to the core, it divides the total payoff equally between the veto players whenever they exist. It is not surprising that the nucleolus does better than the Shapley value at the postulates related to veto players or blockers.

Felsenthal and Machover (1998) introduce the added blocker postulate (ABP). Let v be a simple game, and w another simple game that is obtained by adding an extra player with veto power to v. An index ξ satisfies ABP if whenever a and b are two nondummy players in v we have

$$\frac{\xi_a[w]}{\xi_b[w]} = \frac{\xi_a[v]}{\xi_b[v]}.$$

A flagrant violation of the postulate occurs when $\xi_a[w] > \xi_b[w]$ and $\xi_a[v] < \xi_b[v]$, or the reverse.

Because the nucleolus is in the core, it must give 0 to all players who are not veto players in game w. Thus, the nucleolus violates ABP but not flagrantly. As for the Shapley value, Felsenthal and Machover show that it flagrantly violates ABP.

Another postulate of Felsenthal and Machover is the *blocker share postulate*. This postulate says that, if i is a veto player and S a winning coalition, a P-power index must assign to i at least $\frac{1}{|S|}$. The nucleolus clearly satisfies this postulate, since it divides the payoff equally between all veto players and leaves nothing to outsiders. Felsenthal and Machover show that the Shapley value satisfies this postulate, whereas the Banzhaf, Deegan-Packel and Johnston indices may violate it.

Felsenthal and Machover also point out that a player may lose from becoming a blocker according to the Shapley value. They consider the games $[6;5,3,1,1,1]$ and $[8;5,3,1,1,1]$. The second game is obtained from the first by raising the quota; as a result of this player 1 becomes a veto player. The Shapley value assigns respectively $\left(\frac{3}{5}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}\right)$ and $\left(\frac{11}{20}, \frac{3}{10}, \frac{1}{20}, \frac{1}{20}, \frac{1}{20}\right)$ to these games; the corresponding values for the nucleolus are $\left(\frac{3}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}\right)$ and $(1,0,0,0,0)$. Clearly, if a player is *the only one* to become ablocker he cannot lose because the nucleolus gives him 1; he may lose if other players become blockers as well but this doesn't seem too paradoxical. For example consider the game $[7;6,3,2,1,1]$, whose nucleolus is again $\left(\frac{3}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}\right)$; if the quota is raised to 12 there are three veto players and the new nucleolus is $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0\right)$. Player 1 loses but this is not surprising because he has become symmetric to players 2 and 3, while originally he was more powerful.

Because the nucleolus may give a payoff of 0 to players that are not dummies, it violates one of Felsenthal and Machover (1998) postulates for a power index. It also fails another postulate that Felsenthal and Machover consider essential: it does not respect dominance.

We say that a dominates (or is more desirable than) b if $S \cup \{a\}$ is winning whenever $S \cup \{b\}$ is winning for all S such that $a \notin S$, $b \notin S$. The dominance relation is denoted by $a \succeq b$. We say that a strictly dominates (or is strictly more desirable than) b if $a \succeq b$ but not $b \succeq a$. This is denoted by $a \succ b$. A power measure ξ *respects dominance* if whenever $a \succ b$ in v then $\xi_a[v] > \xi_b[v]$.

The nucleolus satisfies a weak version of this property, namely that $a \succeq b$ implies $\xi_a[v] \geq \xi_b[v]$. However, it is possible for two players to get the same payoff according to the nucleolus even though one of the players is strictly more desirable. The game $[5;3,2,2,1]$ in Example 2 illustrates this possibility: player 1 is more desirable than player 2, but both are given the same payoff by the nucleolus. The list of minimal winning coalitions in this game is $\{1,2\}$, $\{1,3\}$, $\{2,3,4\}$. Players 2 and 3 are less powerful than player 1 because they cannot form a winning coalition on their own, but once we accept that player 4 must get 0 in the nucleolus this is no longer a disadvantage.

The nucleolus can fail to respect dominance even if no player is getting 0, as the following example illustrates.

Example 7 Consider the weighted majority game $[8; 4, 3, 3, 2, 2]$. Players with larger weights are strictly more desirable, but the nucleolus assigns a payoff of $\frac{1}{5}$ to all players.

There are 8 minimal winning coalitions of 4 types in this game: [433], four coalitions of type [432], [422] and two coalitions of type [332]. All of these coalitions have exactly three players.[9]

Suppose all player types have different payoffs. Then coalition types [433] and [432] would be ruled out as too expensive, leaving [422] and [332] . In order for this set to be balanced, both types of coalitions must have a positive probability (otherwise either type [4] or type [3] would have a positive payoff but would not appear in b_1, contradicting Kohlberg's result). But then a player of type [2] appears strictly more often than a player of type [4] and the set cannot be balanced. The same reasoning applies to the case in which two but not all three of the types have the same payoff.

If all players have the same payoff, we can find probabilities for each type of minimal winning coalition: for example we can assign $\frac{2}{15}$ to type [433], $\frac{1}{30}$ to each coalition of type [432], $\frac{1}{3}$ to [422] and $\frac{1}{5}$ to each coalition of type [332] (balancing weights are not unique, but the region of possible balancing weights is quite small in this example). Note that even though we may be forced to assign very different balancing weights to different coalitions in order to achieve a balanced collection, Kohlberg's result ensures that we are never forced to ignore any of the coalitions altogether.

These undesirable properties of the nucleolus do not occur in constant-sum weighted majority games. Peleg (1968) shows that the nucleolus is always a representation of this type of games, and thus must assign a positive payoff to all nondummies and respect dominance.

The idea of Peleg's result is as follows. Take a constant-sum weighted majority game, let w' be an arbitrary vector of weights for the game, and let $q' := \min_{S \in W} w'(S)$ be the quota. The value q' must be strictly greater than 0.5, or the game would not be constant-sum (given a winning coalition S with total weight q', we would find that $w'(N \backslash S) = 1 - q' \geq q'$, contradicting the fact that $N \backslash S$ is a losing coalition). The nucleolus must allocate at least q' to all winning coalitions, so that the maximum excess is at most $1 - q'$ (otherwise there would be a coalition with a greater excess than $1 - q'$, contradicting the definition of the nucleolus). Denote the nucleolus by w and let $q := w(S)$ where S is any of the winning

[9] Individual players with 4, 3 and 2 votes belong to 6, 5 and 4 minimal winning coalitions respectively, hence both the public good index of Holler and Packel (1983) and the Deegan and Packel (1978) index respect dominance in this particular game.

coalitions of maximum excess according to the nucleolus. Then $[q; w]$ is a representation: all winning coalitions are getting at least q, and, since the complement of a losing coalition is a winning coalition and $q \geq q' > \frac{1}{2}$, all losing coalitions are getting less than q.

Peleg's result together with Example 7 show that whether a game is constant-sum or not may make a lot of difference for the nucleolus. Example 7 is as close as possible to being constant-sum in the sense that the quota is no greater than it needs to be in order to keep the game proper. Adding one more player with 1 vote and keeping the quota at 8 would lead to $[8; 4, 3, 3, 2, 2, 1]$, which is a constant-sum game and hence has a nucleolus that respects dominance (the nucleolus of the new game is precisely $\frac{1}{15}(4, 3, 3, 2, 2, 1)$).

The nucleolus may still be a representation if the game is not constant sum. For homogeneous games, Peleg and Rosenmüller (1992) show that the nucleolus is a representation if the set of minimal winning coalitions is balanced. The nucleolus may still be a representation if the game is neither homogeneous nor constant-sum; an example is $[6; 3, 2, 2, 1, 1]$.[10]

Straffin (1998) points out that the Banzhaf index and the Shapley value may rank players differently in the game $[2; 1111] \otimes [3; 2111]$, where the notation $\otimes$ means that a majority must be obtained in both voting bodies. This example also shows that the Banzhaf index and the nucleolus may also rank players differently; it also shows that the nucleolus may appear counterintuitive as a measure of power: it assigns 0 to all players in the first game, and $\left(\frac{2}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}\right)$ to the players in the second game. In contrast, the Banzhaf index assigns more to players of the first type than to players of the third type. The nucleolus must assign 0 value to the players of the first type because otherwise the set of coalitions of maximum excess would not be balanced. Thus, whether we take the Shapley value or the nucleolus as a measure of P-power, a player can be more powerful than another under office-seeking behavior, but less powerful under policy-seeking behavior. Felsenthal and Machover refer to this possibility as “somewhat paradoxical”.

6 Concluding Remarks

This chapter makes a case for the nucleolus as a power index in divide-the-dollar games, especially if the nucleolus is a representation of the game. The nucleolus can be interpreted as a competitive price system and has relatively solid noncooperative foundations. At a more fundamental level, the nucleolus identifies a set of attractive coalitions, whereas the Shapley value is determined by all coalitions.

It is common wisdom in the power indices literature that ”the very idea behind voting power is that the weight of a voter is not a good measure of power” (Pajala

[10] Looking at the early stages of the EEC Council of Ministers, the nucleolus is a representation for the 1958 and the 1995 Councils but not for 1973, 1981 or 1986 (see Le Breton et al 2012).

2002). If we adopt the nucleolus as a power index, weights will be power for some games (including all constant-sum weighted majority games), provided that we choose the right weights to represent the game. Interestingly, homogeneity is neither necessary nor sufficient for weights to be a measure of power.

The Deegan-Packel index assume that only minimal winning coalitions will form, each of them with equal probability, and players will divide the payoff equally. Clearly, the nucleolus does not assume that coalitions divide the payoff equally. It does not assume either that all minimal winning coalitions form. Some minimal winning coalitions may not be of maximum excess, like a coalition of type [222] in Example 1. On the other hand, some coalitions of maximum excess may not be minimal winning, like a coalition of type [321] in Example 2. The nucleolus does not assume equiprobability of coalitions, but it does imply equiprobability of players (at least, of the players that get a positive payoff), which is an appealing property.

References

Albers, W. (1974). Core- and kernel-variants based on imputations and demand profiles. In O. Moeschlin & D. Palaschke (Eds.), *Game Theory and Related Topics*. Amsterdam:North-Holland.

Banks, J. S., & Duggan, J. (2000). A bargaining model of collective choice. *American Political Science Review*, *94*, 73–88.

Banzhaf, J. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review*, *19*, 317–343.

Baron, D. P., & Ferejohn, J. A. (1989). Bargaining in legislatures. *American Political Science Review*, *83*, 1181–1206.

Bennett, E. (1983). The aspiration approach to predicting coalition formation and payoff distribution in sidepayment games. *International Journal of Game Theory*, *12*, 1–28.

Binmore, K. (1987). Perfect equilibria in bargaining models. In K. Binmore & P. Dasgupta (Eds.), *The Economics of Bargaining*. Oxford: Blackwell.

Cross, J. (1967). Some theoretical characteristics of economic and political coalitions. *Journal of Conflict Resolution*, *11*, 184–195.

Davis, M., & Maschler, M. (1965). The kernel of a cooperative game. *Naval Research Logistics Quarterly*, *12*, 223–259.

Davis, M., & Maschler, M. (1967). Existence of stable payoff configurations for cooperative games. In M. Shubik (Ed.), *Essays in Mathematical Economics in Honor of Oskar Morgenstern*: Princeton University Press.

Deegan, J., & Packel, E. W. (1978). A new index of power for simple n-person games. *International Journal of Game Theory*, *7*, 113–123.

Derks, J., & Kuipers, J. (1997). Implementing the simplex method for computing the prenucleolus of transferable utility games: Unpublished manuscript.

Dubey, P. (1975). On the uniqueness of the Shapley value. *International Journal of Game Theory*, *4*, 131–139.

Eraslan, H., & McLennan, A. (2006). Typescript: Uniqueness of stationary equilibrium payoffs in coalitional Bargaining.

Felsenthal, D. S., & Machover, M. (1998). *The measurement of voting power: Theory and practice, problems and paradoxes*. Camberley: Edward Elgar Publishing.

Gul, F. (1989). Bargaining foundations of the Shapley value. *Econometrica*, *57*, 81–96.

Harrington, J. (1990). The power of the proposal maker in a model of endogenous agenda formation. *Public Choice*, *64*, 1–20.

Hart, S., & Mas-Colell, A. (1996). Bargaining and value. *Econometrica*, *64*, 357–380.

Holler, M. J., & Packel, E. W. (1983). Power, luck and the right index. *Journal of Economics*, *43*, 21–29.

Johnston, R. J. (1978). On the measurement of power: Some reactions to Laver. *Environment and Planning A*, *10*, 907–914.

Kalandrakis, T. (2006). Proposal rights and political power. *American Journal of Political Science*, *50*, 441–448.

Kohlberg, E. (1971). On the nucleolus of a characteristic function game. *SIAM Journal of Applied Mathematics*, *20*, 62–66.

Laruelle, A., & Valenciano, F. (2001). Shapley-Shubik and Banzhaf indices revisited. *Mathematics of Operations Research*, *26*, 89–104.

Laruelle, A., & Valenciano, F. (2007). Bargaining in committees as an extension of Nash's bargaining theory. *Journal of Economic Theory*, *132*, 291–305.

Laruelle, A., & Valenciano, F. (2008). Noncooperative foundations of bargaining power in committees and the Shapley-Shubik index. *Games and Economic Behavior*, *63*, 341–353.

Le Breton, M., & Zaporozhets, V. (2010). Sequential legislative lobbying under political certainty. *Economic Journal*, *120*, 281–312.

Le Breton, M., Montero, M., & Zaporozhets, V. (2012). Voting power in the EU Council of Ministers and fair decision making in distributive politics, forthcoming in Mathematical Social Sciences.

Maschler, M. (1992). The bargaining set, kernel, and nucleolus. In R. J. Aumann & S. Hart (Eds.), *Handbook of Game Theory with Economic Applications* (Vol. I). Amsterdam: North-Holland.

Montero, M. (2001). The nucleolus as a consistent power index in noncooperative majority games, CentER Discussion Paper No 2001–39, later published in. *Games and Economic Behavior*, *54*, 380–397.

Montero, M. (2008), Proportional payoffs in majority games, Available at SSRN: http://ssrn.com/abstract=1112626

Napel, S., & Widgrén, M. (2001). Inferior players in simple games. *International Journal of Game Theory*, *30*, 209–220.

Nohn, A. (2010a). *Veto players and non-cooperative foundations of power in coalitional bargaining*. Typescript.

Nohn, A. (2010b). *Coalitional bargaining with Markovian proposers*. Typescript.

Pajala, A. (2002), Voting Power and Power Index Website: a Voting Power WWW-Resource Including POWERSLAVE Voting Body Analyser [online], University of Turku, http://powerslave.val.utu.fi

Peleg, B. (1968). On weights of constant-sum majority games. *SIAM Journal of Applied Mathematics*, *16*, 527–532.

Peleg, B., & Rosenmüller, J. (1992). The least-core, nucleolus and kernel of homogeneous weighted majority games. *Games and Economic Behavior*, *4*, 588–605.

Pérez-Castrillo, D., & Wettstein, D. (2001). Bidding for the surplus: A non-cooperative approach to the Shapley value. *Journal of Economic Theory*, *100*, 274–294.

Roth, A. (1977a). The Shapley value as a von Neumann-Morgenstern utility. *Econometrica*, *45*, 657–664.

Roth, A. (1977b). Utility functions for simple games. *Journal of Economic Theory*, *16*, 481–489.

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica*, *50*, 97–108.

Schmeidler, D. (1969). The nucleolus of a characteristic function game. *SIAM Journal on Applied Mathematics*, *17*, 1163–1170.

Shapley, L. S. (1953). A value for n-person games. *Annals of Mathematical Studies*, *28*, 307–317.

Shapley, L. S., & Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review*, *81*, 935–945.

Straffin, P. D. (1998). The Shapley-Shubik and Banzhaf power indices as probabilities. In A. Roth (Ed.), *The Shapley Value*. Cambridge: Cambridge University Press.

Sudhölter, P. (1996). The modified nucleolus as canonical representation of weighted majority games. *Mathematics of Operations Research*, *21*, 734–756.

Vidal-Puga, J. J. (2008). Forming coalitions and the Shapley NTU value. *European Journal of Operational Research*, *190*, 659–671.

Winter, E. (1994). The demand commitment bargaining and snowballing cooperation. *Economic Theory*, *4*, 255–273.

Winter, E. (1996). Voting and vetoing. *American Political Science Review*, *90*, 813–823.

Young, H. P. (1978). A tactical lobbying game. In P. C. Ordeshook (Ed.), *Game theory and political science*. New York: New York University Press.

Coalition Configurations and the Public Good Index

José M. Alonso-Meijide, Balbina Casas-Méndez, M. Gloria Fiestras-Janeiro, Manfred J. Holler and Andreas Nohn

1 Introduction

Power indices are quantitative measures to express power in simple games. The most important power indices are: the Shapley-Shubik index (Shapley and Shubik 1954), the Banzhaf-Coleman index (Banzhaf 1965; Coleman 1971), the Deegan-Packel index (Deegan and Packel 1978), the Public Good Index (Holler 1982), and the Johnston power index (Johnston 1978).

There exists a vast literature on modifications of these original power indices. Most of these variations are proposed to analyze situations where information about

This chapter has been published in Homo Oeconomicus 29 (3), 2012.

J. M. Alonso-Meijide (✉) · B. Casas-Méndez
Department of Statistics and Operations Research, University of Santiago de Compostela, Santiago de Compostela, Spain
e-mail: josemaria.alonso@usc.es

B. Casas-Méndez
e-mail: balbina.casas.mendez@usc.es

M. G. Fiestras-Janeiro
Department of Statistics and Operations Research, University of Vigo, Vigo, Spain
e-mail: fiestras@uvigo.es

M. J. Holler
Institute of SocioEconomics, University of Hamburg, Germany and Public Choice Research Centre, Turku, Finland
e-mail: holler@econ.uni-hamburg.de

A. Nohn
Public Choice Research Centre, Turku, Finland
e-mail: nohn@econ.uni-hamburg.de

A. Nohn
Institute of SocioEconomics, University of Hamburg, Hamburg, Germany

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*, DOI: 10.1007/978-3-642-35929-3_16, © Springer-Verlag Berlin Heidelberg 2013

the behavior of players is available and, e.g., a priori unions of players exist and are known. The most studied modification of the traditional power indices is the Owen value (Owen 1977). We name it the Owen index. The Owen value is a two-step extension of the Shapley value (Shapley 1953) that takes a priori unions (or a coalition structure) into consideration. In the first step, a game among unions (the quotient game) is considered. The Owen index satisfies the quotient game property, that is, the sum of the power assigned to the players of a union coincides with the power of the union in the quotient game. Another interesting property of the Owen index is symmetry in the quotient game: given two unions which play symmetric roles in the quotient game, they are awarded with the same apportionment of the power. The Owen index, of course, coincides with the Shapley-Shubik index if each a priori union contains one element only. It is well known that the Shapley-Subik index refers to permutations of players when modelling coalition formation, taking into account all winning coalitions. The Public Good Index, which we extend in the following, considers minimal winning coalitions only.

In this chapter, we consider two extensions of the Public Good Index for games with a priori unions: the Union Public Good Index (Holler and Nohn 2009) and the Solidarity Public Good Index (Alonso-Meijide et al. 2010a). The Union Public Good Index follows the main argument of the original Public Good Index, translated to the model with a priori unions, that is, only minimal winning coalitions of the quotient game are relevant. The Solidarity Public Good Index stresses the public good property which suggests that all members of a winning coalition derive equal power, irrespective of their possibility to form alternative coalitions. The Solidarity Public Good Index satisfies the properties of quotient game and symmetry in the quotient game. The Union Public Good Index does not satisfy these properties but assigns power in a different way.

The coalition configurations are extensions of the coalition structures and allow to represent more complex relations among players. A coalition configuration is defined by a family of coalitions whose union is the whole set of players but a player can belong to several coalitions. A player can agree on some interests with other players, if we consider, for example, ideological reasons, but he can be far from these players if we consider economic interests. In this case, players do not form necessarily a partition of the set of players, because each player can belong to one or several coalitions. The model of coalition configuration was considered in Albizuri et al. (2006) where a generalization of the Owen index (the configuration value) was proposed. In Albizuri and Aurrekoetxea (2006) the Banzhaf-Coleman index was generalized for the family of games with a coalition configuration. More recently, Andjiga and Courtin (2010) developed three indices for games with coalition configurations, namely Owen-Shapley-CCF share, Owen-Banzhaf-CCF share, and Deegan-Packel-CCF share, by using the concept of share function as introduced by van der Lann and van den Brink (2002, 2005).

In what follows we apply the concept of coalition configuration to the Public Good Index and present two generalizations. These two variations are based on the Union Public Good Index and the Solidarity Public Good Index. For these extensions we provide axiomatic characterizations. A real-world example illustrates the new indices.

The paper is organized as follows. In Sect. 2, some basic definitions and characterizations of different variants of the Public Good Index for games with a priori unions are introduced. In Sect. 3, we present two variants of the Public Good Index for games with coalition configurations and give axiomatic characterizations. Finally, in Sect. 4, we illustrate and compare these generalizations evaluating the previous indices in the Catalonian Parliament held in 2006 elections.

2 Preliminaries

2.1 Simple Games and the Public Good Index

A *simple game* is a pair (N, v) where $N = \{1, \ldots, n\}$ is a finite *set of players* and v is the *characteristic function* that allocates to each coalition $S \subseteq N$ the value 0 or 1 in such a way that $v(\emptyset) = 0$, $v(N) = 1$, and $v(S) = 1$ implies that $v(T) = 1$ for all $S \subseteq T$. We say that $S \subseteq N$ is a *minimal winning coalition* if $v(S) = 1$ and $v(T) = 0$ for all $T \not\subseteq S$. $W(v)$ denotes the *set of minimal winning coalitions* of (N, v).

A *power index* is a mapping f assigning each simple game (N, v) an n-dimensional real valued vector $f(N, v) = (f_1(N, v), \ldots, f_n(N, v))$. Based on the assumptions that coalitional values are public goods and only minimal winning coalitions are relevant when it comes to power, the *Public Good Index* (PGI) proposed by Holler (1982) assigns power in a proportional way to the number of minimal winning coalitions a player belongs to. Denoting $W_i(v)$ as the set of minimal winning coalitions containing i, the PGI δ is given by

$$\delta_i(N, v) = \frac{|W_i(v)|}{\sum_j |W_j(v)|}, \ i = 1, \ldots, n.$$

Holler and Packel (1983) characterize the PGI as the unique power index satisfying efficiency, symmetry, null player, and PGI-mergeability. An index f satisfies *efficiency* if $\sum_i f_i(N, v) = 1$ for all simple games (N, v). Two players i and j are *symmetric* if $v(S \cup \{i\}) = 1$ if and only if $v(S \cup \{j\}) = 1$ for all $S \subseteq N \backslash \{i, j\}$. A player i is called a *null player* if $S \backslash \{i\} \in W(v)$ for all coalitions $S \in W(v)$. A power index f satisfies *symmetry* if $f_i(N, v) = f_j(N, v)$ for all symmetric players i and j. A power index f satisfies *null player* if $f_i(N, v) = 0$ for all null players i. Two simple games (N, v) and (N, v') are *mergeable* if $S \in W(v)$ implies $v'(S) = 0$ and $S \in W(v')$ implies $v(S) = 0$. In particular, the sets of minimal winning coalitions $W(v)$ and $W(v')$ are disjoint. The *merged game* $(N, v \oplus v')$ of two mergeable games (N, v) and (N, v') is the simple game such that $W(v \oplus v') = W(v) \cup W(v')$. Now, a power index f satisfies *PGI-mergeability* if for all mergeable games (N, v) and (N, v'),

$$f(N, v \oplus v') = \frac{\sum_i |W_i(v)| f(N, v) + \sum_i |W_i(v')| f(N, v')}{\sum_i (|W_i(v)| + |W_i(v')|)}.$$

2.2 Simple Games with a Priori Unions and Public Good Indices

For a set of players N, a *set of a priori unions* is a partition $P = \{P_1, \ldots, P_m\}$ of N, that is, a family of nonempty and mutually disjoint subsets of N whose union coincides with N. We denote by P^n the partition where each player forms his own union, that is, $P^n = \{\{i\} | i \in N\}$. We also use P as the mapping assigning each player i the *union* $P(i) \in P$ he is a member of. A *simple game with a priori unions* is a triplet (N, v, P), that is, a set of players N, a characteristic function v, and a set of a priori unions P on N.

Given (N, v, P), the corresponding *quotient game* is the simple game (M, v^P) with player set $M = \{1, \ldots, m\}$ and characteristic function v^P, where $v^P(R) = v(\cup_{k \in R} P_k)$, for each $R \subseteq M$. A coalition $R \subseteq M$ in the quotient game is winning if and only if the coalition of represented unions $\bigcup_{k \in R} P_k$ is winning in (N, v). We denote the set of minimal winning coalitions in the quotient game by $W(v^P)$ and by $W_k(v^P)$ the set of minimal winning coalitions containing union $k \in M$.

A union P_k's power in the quotient game, measured by the PGI, amounts to

$$\delta_k(M, v^P) = \frac{|W_k(v^P)|}{\sum_k |W_k(v^P)|}, \quad k = 1, \ldots, m.$$

A *coalitional power index* is a mapping f assigning each simple game with a priori unions (N, v, P) an n-dimensional real valued vector $f(N, v, P) = (f_1(N, v, P), \ldots, f_n(N, v, P))$.

2.2.1 Coalitional PGIs and Quotient Game Property

Given a power index g, a coalitional power index f is a *coalitional g -power index* if for every simple game (N, v) it holds $f(N, v, P^n) = g(N, v)$, that is, if f is equal to g in case each a priori union is a singleton coalition. Then, a coalitional power index f is a coalitional Public Good Index, if $f(N, v, P^n) = \delta(N, v)$ for every simple game (N, v).

The following four properties constitute the analogues of the four axioms of the PGI, stated in Sect. 2.1, for coalitional power indices if the a priori unions are singletons and P^n applies. A coalitional power index f satisfies

- *singleton efficiency* if for every simple game (N, v), $\sum_i f_i(N, v, P^n) = 1$.
- *singleton null player* if for every simple game (N, v), $f_i(N, v, P^n) = 0$ for every null player i in (N, v).

- *singleton symmetry* if for every simple game (N, v), $f_i(N, v, P^n) = f_j(N, v, P^n)$ for every symmetric players i and j in (N, v).
- *singleton PGI-mergeability* if for every two mergeable simple games (N, v) and (N, v'),

$$f(N, v \oplus v', P^n) = \frac{\sum_i |W_i(v)| f(N, v, P^n) + \sum_i |W_i(v')| f(N, v', P^n)}{\sum_i (|W_i(v)| + |W_i(v')|)}.$$

Proposition 1 *A coalitional power index f is a coalitional Public Good Index if and only if it satisfies singleton efficiency, singleton null player, singleton symmetry, and singleton PGI-mergeability.*

In the context of coalitional power indices, we can consider the next property that gives an interesting relation between the power of players of a priori union and the power of this union in the quotient game. A coalitional power index f satisfies *quotient game property* if for every simple game with a priori unions (N, v, P) and every union $P_k \in P$,

$$\sum_{i \in P_k} f_i(N, v, P) = f_k(M, v^P, P^m).$$

This property was proposed by Winter (1992) to characterize the Owen value in the more general model of TU games with a priori unions. In the model of simple games with a priori unions, the quotient game property states that the sum of the power obtained by the players of a union coincides with the power obtained by this union in the game played by the unions, i.e., the quotient game. In the presence of the quotient game property, three of the singleton properties are significantly enhanced. Singleton efficiency implies

- *efficiency*: for every (N, v, P), $\sum_{i \in N} f_i(N, v, P) = 1$.

Singleton null player implies

- *null union*: for every (N, v, P) and every union $P_k \in P$, such that k is a null player in the quotient game (M, v^P), $\sum_{i \in P_k} f_i(N, v, P) = 0$.

It does *not*, however, also imply

- *null player*: for every (N, v, P) and every null player i in (N, v), $f_i(N, v, P) = 0$.

In fact, there are coalitional indices that satisfy the quotient game property and singleton null player (and, then, null union), but not null player. For instance, the Union PGI and the Solidarity PGI that we introduce below. Besides, singleton symmetry implies

- *symmetry among unions*: for every (N, v, P) and all unions $P_k, P_{k'}$ which are symmetric players in the quotient game (M, v^P), $\sum_{i \in P_k} f_i(N, v, P) = \sum_{i \in P_{k'}} f_i(N, v, P)$.

In analogy to singleton null player which does not imply null player, singleton symmetry does *not* imply

- *symmetry within unions*: for every (N, v, P) and all symmetric players $i, j \in P_k$ in (N, v), $f_i(N, v, P) = f_j(N, v, P)$.

Also, note that quotient game property does not necessarily extend singleton PGI-mergeability to any stronger version of PGI-mergeability.

For coalitional Public Good Indices which satisfy the quotient game property also satisfy efficiency, null union, and symmetry among unions while they do not necessarily satisfy null player, symmetry within unions or any particular form of PGI mergeability. One obtains the following identity stating a union P_k's overall power is equal to its Public Good Index in the quotient game,

$$\sum_{i \in P_k} f_i(N, v, P) = \delta_k(M, v^P).$$

2.2.2 Axiomatizations of PGIs for A Priori Unions

Alonso-Meijide et al. (2010a) introduce and axiomatize the *Solidarity Public Good Index*. This index distributes power in two steps. In the first step, it assigns power to each union equal to its PGI in the quotient game (thus satisfying quotient game property). In the second step, the Solidarity Public Good Index Υ stresses the public good property by assigning equal power to each member of the same a priori union,

$$\Upsilon_i(N, v, P) = \delta_{P(i)}(M, v^P) \frac{1}{|P(i)|}, \quad i = 1, \ldots, n.$$

It thus satisfies

- *solidarity*: for every (N, v, P), $f_i(N, v, P) = f_j(N, v, P)$ for all players i, j being member of the same union, that is, $P(i) = P(j)$.

Alonso-Meijide et al. (2010a) provide an axiomatization of the Solidarity PGI using, among others, the following two properties.

- *Independence of superfluous coalitions*: for every (N, v, P) and (N, v', P) with $W(v^P) = W(v'^P)$, $f(N, v, P) = f(N, v', P)$.
- *PGI-mergeability in the quotient game*: for every (N, v, P) and (N, v', P) where the quotient games (M, v^P) and (M, v'^P) are mergeable,

$$f(N, v'', P) = \frac{\sum_k |W_k(v^P)| f(N, v, P) + \sum_k |W_k(v'^P)| f(N, v', P)}{\sum_k (|W_k(v^P)| + |W_k(v'^P)|)},$$

for every simple game (N, v''), with characteristic function v'' such that $W(v'') \subseteq W(v) \cup W(v')$ and $W(v''^P) = W(v^P) \cup W(v'^P)$.

Proposition 2 Alonso-Meijide et al. (2010a) *The Solidarity PGI* Υ *is the unique coalitional power index satisfying efficiency, null union, symmetry among unions, solidarity, independence of superfluous coalitions, and PGI-mergeability in the quotient game.*

There is, however, an obvious alternative referring to the quotient game property.

Proposition 3 (Alonso-Meijide et al. 2010b) *The Solidarity PGI* Υ *is the unique coalitional PGI satisfying quotient game property and solidarity.*

Holler and Nohn (2009) introduce four variants of the PGI for a priori unions. The first variant, the *Union Public Good Index* Λ, is as close as possible to the original spirit of the PGI. It is based on the assumptions that the coalitional value is a public good and only minimal winning coalitions are relevant. The second assumption applies to coalitions being minimal with respect to the a priori unions and the simple game that has the a priori unions as players. The player's power is, then, proportional to the number of minimal winning coalitions his union is a member of in the quotient game, that is,

$$\Lambda_i(N, v, P) = \frac{|W_{P(i)}(v^P)|}{\sum_k |P_k||W_{P_k}(v^P)|}, \quad i = 1, \ldots, n.$$

As with the Solidarity PGI, all members of the same union have equal power. However, the Union PGI is the only of the overall extensions not assigning power to unions on the basis of the PGI in the corresponding quotient game. We hence provide an axiomatization not directly using its being a coalitional PGI but more elementary axioms. For this sake, we say that a coalitional power index f satisfies

- *symmetry among players of symmetric unions*: for every (N, v, P), and for all members i, j of symmetric unions $P(i)$ and $P(j)$, $f_i(N, v, P) = f_j(N, v, P)$.
- *PGI-mergeability among unions*: for every (N, v, P) and (N, v', P) where the quotient games (M, v^P) and (M, v'^P) are mergeable,

$$f(N, v'', P) = \frac{\sum_k |P_k||W_k(v^P)|f(N, v, P) + \sum_k |P_k||W_k(v'^P)|f(N, v', P)}{\sum_k |P_k|(|W_k(v^P)| + |W_k(v'^P)|)}$$

  for every simple game (N, v''), with characteristic function v'' such that $W(v'') \subseteq W(v) \cup W(v')$ and $W(v''^P) = W(v^P) \cup W(v'^P)$.

Proposition 4 (Alonso-Meijide et al. 2010b) *The Union Public Good Index* Λ *is the unique coalitional power index satisfying efficiency, null union, symmetry among players of symmetric unions, independence of superfluous coalitions, and PGI-mergeability among unions.*

3 PGIs for Coalition Configurations

For a set of players N, a *coalition configuration* is a finite collection of nonempty subsets of $N, C = \{C_1, \ldots, C_m\}$, with the only assumption that each player belongs to at least one subset, that is, $\cup_{k \in M} C_k = N$ where $M = \{1, \ldots, m\}$. We explicitly allow for non-disjoint subsets, i.e., that $C_k \cap C_{k'} \neq \emptyset$ for some $k, k' \in M$. Elements of C are called coalitions.

A family of a priori unions is a particular case of a coalition configuration. Our aim is to extend the Solidarity Public Good Index, Υ, and the Union Public Good Index, Λ, to this new model.

We denote by (N, v, C) a *simple game with a coalition configuration* where (N, v) is a simple game and C a coalition configuration. Given a player $i \in N$, $C(i)$ denotes the elements of C containing i, that is, $C(i) = \{C_k \in C \,:\, i \in C_k\}$. (M, v^C) denotes the game played among the elements of C, that is, $v^C(R) = v\big(\cup_{j \in R} C_j\big)$, for every $R \subseteq M$. $W(v^C)$ denotes the set of minimal winning coalitions of this game and $W_k(v^C)$ the subset of minimal coalitions of this game containing $k \in M$.

A *coalition configuration power index* is a mapping f assigning each simple game with a coalition configuration (N, v, C) an n-dimensional real valued vector $f(N, v, C) = (f_1(N, v, C), \ldots, f_n(N, v, C))$.

We introduce two coalition configuration power indices.

Definition 5

1. The Generalized Solidarity PGI is given by

$$\Upsilon_i(N, v, C) = \frac{1}{\sum_{l \in M} |W_l(v^C)|} \sum_{C_k \in C(i)} \frac{\big|W_k(v^C)\big|}{|C_k|}, \quad i = 1, \ldots, n.$$

2. The Generalized Union PGI is given by

$$\Lambda_i(N, v, C) = \frac{1}{\sum_{l \in M} |W_l(v^C)||C_l|} \sum_{C_k \in C(i)} \big|W_k\big(v^C\big)\big|, \; i = 1, \ldots, n.$$

Example 6 Take (N, v, C) such that $N = \{1, 2, 3, 4, 5\}$, $W(v) = \{\{1, 2, 3\}, \{1, 2, 4, 5\}\}$ and $C = \{C_1, C_2, C_3\}$, where $C_1 = \{1, 2\}, C_2 = \{1, 3\}$, and $C_3 = \{3, 4, 5\}$. Then $C(1) = \{C_1, C_2\}, C(2) = \{C_1\}, C(3) = \{C_2, C_3\}$ and $C(4) = C(5) = \{C_3\}$. The game played by the coalitions is given by (M, v^C) where $M = \{1, 2, 3\}$ and $W(v^C) = \{\{C_1, C_2\}, \{C_1, C_3\}\}$. Then, $|W_1(v^C)| = 2$, $|W_2(v^C)| = 1$, and $|W_3(v^C)| = 1$. For player 1, we have

$$\Upsilon_1(N, v, C) = \frac{1}{4} \times \left(\frac{2}{2} + \frac{1}{2}\right) = \frac{3}{8}$$

and

$$\Lambda_1(N,v,C) = \frac{2+1}{2\times 2 + 1\times 2 + 1\times 3} = \frac{3}{9}.$$

Following a similar procedure for players 2 to 5, we have

$$\Upsilon(N,v,C) = (3/8, 2/8, 5/24, 1/12, 1/12),$$
$$\Lambda(N,v,C) = (3/9, 2/9, 2/9, 1/9, 1/9).$$

Given a coalitional power index g, a coalitional configuration power index f is a *configuration g-index* if for every simple game with a priori unions (N,v,P) it holds $f(N,v,P) = g(N,v,P)$, that is, if f is equal to g in case of any simple game with a priori unions (N,v,P).

In order to provide a characterization of the above indices we use the concept of configuration g-index and a new property. To introduce this property for coalitional configuration power indices we need some additional notation. Take a simple game with coalition configuration (N,v,C) and consider $i \in N$. The subset of coalition indices to which i belongs to is $M(i) = \{r \in M : i \in C_r\}$. Now we introduce a simple game with a priori unions (N', v'^C, P^C), that we name the *replica game* of (N,v,C). The set of agents of the replica game is $N' = R_1 \cup \cdots \cup R_n = \{r_{ik}\}_{i\in N,\, k\in M(i)}$ where $R_i = \{r_{ik}\}_{k\in M(i)}$ for every $i \in N$. The characteristic function of the replica game is given by $v'^C(S) = v(R_S)$ for every $S \subseteq N'$, where $R_S = \{i \in N : R_i \cap S \neq \emptyset\}$. Finally, the set of a priori unions is $P^C = \{P_1^C, \ldots, P_m^C\}$ where $P_k^C = \{r_{ik}\}_{i\in C_k}$.

The interpretation of the replica game associated with the game with coalition configuration is intuitive. Each player in the original set is replicated in as many new players as the number of coalitions to which it belongs. Each coalition of new players creates a coalition in the initial game if each player is replaced by a replica of him. Thus, the value of a coalition in the new game is defined as the value of the corresponding coalition in the original game. Finally, the partition of the new set of players is constructed as follows: in each coalition of the initial configuration each player is replaced by a replica of himself that has yet not been considered. In accordance with this, we say that *agent $l \in N'$ is a replica of agent $i \in C_k$* if and only if $l \in R_i$ and $l \in P_k^C$. We can now state the property.

- *Addition of replicas*—for every simple game with a coalition configuration (N,v,C) and its replica game (N', v'^C, P^C) it is

$$f_i(N,v,C) = \sum_{l\in R_i} f_l(N', v'^C, P^C),\ \ i = 1,\ldots,n$$

  where R_i is the set of agent i's replicas.

We propose the following characterizations of the coalition configuration power indices introduced in Definition 5. These characterizations are analogous to the chracterizations presented in Proposition 3 and Proposition 4.

Proposition 7 *The Generalized Solidarity PGI*, Υ, *is the unique coalition configuration Solidarity PGI-value satisfying the addition of replicas property.*

Proof First, we prove that the Generalized Solidarity PGI satisfies the addition of replicas property. Take a simple game with a coalition configuration (N, v, C), its replica game (N', v'^C, P^C) and $i \in N$. Then we have:

$$\begin{aligned}\Upsilon_i(N, v, C) \quad &= \frac{1}{\sum_{l \in M} |W_l(v^C)|} \sum_{C_k \in C(i)} \frac{|W_k(v^C)|}{|C_k|} \\ &= \frac{1}{\sum_{l \in M} |W_l(v'^{P^C})|} \sum_{s \in R_i} \frac{|W_{P^C(s)}(v'^{P^C})|}{|P^C(s)|} \\ &= \sum_{s \in R_i} \Upsilon_s(N', v'^C, P^C).\end{aligned}$$

Now we prove the uniqueness part of the result. Assume that f^1 and f^2 are two different coalition configuration Solidarity PGI-values satisfying addition of replicas. Then, there is a coalition configuration simple game (N, v, C) and $i \in N$ such that $f_i^1(N, v, C) \neq f_i^2(N, v, C)$. Since f^1 and f^2 satisfy addition of replicas, then

$$f_i^1(N, v, C) = \sum_{l \in R_i} f_i^1(N', v'^C, P^C) \quad \text{and}$$

$$f_i^2(N, v, C) = \sum_{l \in R_i} f_i^2(N', v'^C, P^C),$$

where (N', v'^C, P^C) is the replica game of (N, v, C). Notice that P^C is a partition of N' and since f^1 and f^2 are coalition configuration Solidarity PGI-values they coincide when we have a partition. Then, $f_i^1(N, v, C) = f_i^2(N, v, C)$, being a contradiction. □

Proposition 8 *The Generalized Union PGI*, Λ, *is the unique coalition configuration Union PGI-value satisfying the addition of replicas property.*

Proof We only prove that the Generalized Union PGI satisfies the addition of replicas property. The uniqueness part follows a similar reasoning that we use in Proposition 7.

Take a simple game with a coalition configuration (N, v, C), its replica game (N', v'^C, P^C) and $i \in N$. Then we have:

$$\begin{aligned}\Lambda_i(N, v, C) &= \frac{1}{\sum_{l \in M} |W_l(v^C)||C_l|} \sum_{C_k \in C(i)} |W_k(v^C)| \\ &= \frac{1}{\sum_{l \in M} |W_l(v'^{P^C})||P_l^C|} \sum_{s \in R_i} |W_{P^C(s)}(v'^{P^C})| \\ &= \sum_{s \in R_i} \Upsilon_s(N', v'^C).\end{aligned}$$

□

4 An Application

Alonso-Meijide et al. (2010a) analyze the Parliament of Catalonia held from elections in 2006 by computing two extensions of the PGI index to the context of simple games with a priori unions. Besides, they consider two different families of a priori unions in an independent way. These families arise from assuming similarities among parties in accordance with two criteria: the independence dimension from Spanish centralism versus Catalanism, on the one hand, and, on the other hand, the ideological dimension of left and right. For the sake of clarity we repeat the data of this example.

The Parliament of Catalonia consists of 135 members. Following these elections, the Parliament was composed of:

1. 48 members of *CIU*, *Convergéncia i Unió*, a Catalan nationalist middle-of-the-road party,
2. 37 members of *PSC*, *Partido de los Socialistas de Cataluña*, a moderate left-wing socialist party federated to the *Partido Socialista Obrero Español*,
3. 21 members of *ERC*, *Esquerra Republicana de Cataluña*, a radical Catalan nationalist left-wing party,
4. 14 members of *PPC*, *Partido Popular de Cataluña*, a conservative party which is a Catalan delegation of the *Partido Popular*,
5. 12 members of *ICV*, *Iniciativa por Cataluña-Los Verdes-Izquierda Alternativa*, a coalition of ecologist groups and Catalan eurocommunist parties federated to *Izquierda Unida*, and
6. 3 members of *CPC*, *Ciudadanos-Partidos de la Ciudadanía*, a non-Catalanist party.

We identify *CIU* as player 1, *PSC* as player 2 , *ERC* as player 3, *PPC* as player 4, *ICV* as player 5 and *CPC* as player 6. Then, taking $N = \{1,2,3,4,5,6\}$ as the set of players, the corresponding set of minimal winning coalitions is

$$W(v) = \{\{1,2\},\{1,3\},\{1,4,5\},\{2,3,4\},\{2,3,5\}\}.$$

We see that *CPC* is a null player. We consider two possible partitions of players according to the independence dimension and the ideological dimension, respectively,

$$P^1 = \{\{1\},\{2\},\{3,5\},\{4\},\{6\}\} \text{ and } P^2 = \{\{1\},\{2,3,5\},\{4\},\{6\}\},$$

where P^1 represents the dimension of Spanish centralism versus Catalanism while P^2 represents the a priori unions that correspond to the left-right dimension.

The model of coalition configuration allows us to consider several criteria simultaneously. In our example we simultaneously take into account both criteria, i.e., the coalition configuration given by

$$C = \{C_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8, C_9\}$$

Table 1 Several indices

Party	$\Upsilon_i(P^1)$	$\Upsilon_i(P^2)$	$\Upsilon_i(C)$	$\Lambda_i(P^1)$	$\Lambda_i(P^2)$	$\Lambda_i(C)$
CIU	0.3333	0	0.3636	0.25	0	0.25
PSC	0.3333	0.3333	0.3030	0.25	0.3333	0.25
ERC	0.1667	0.3333	0.1667	0.25	0.3333	0.25
PPC	0	0	0	0	0	0
ICV	0.1667	0.3333	0.1667	0.25	0.3333	0.25
C's	0	0	0	0	0	0

where

$$C_1 = \{1\},\ C_2 = \{2\},\ C_3 = \{3,5\},\ C_4 = \{4\},\ C_5 = \{6\},$$
$$C_6 = \{1\},\ C_7 = \{2,3,5\},\ C_8 = \{4\},\ C_9 = \{6\}.$$

The minimal winning coalitions of the game (M, v^C) are the following

$$\{C_7\},\ \{C_1, C_2\},\ \{C_1, C_3\},\ \{C_2, C_3\},\ \{C_2, C_6\},\ \{C_3, C_6\}.$$

Table 1 presents the values of the Solidarity PGI and Union PGI computed using the a priori unions P_1 and P_2. It also shows the values of the Generalized Solidarity PGI and the Generalized Union PGI related to the coalition configuration C.

Notice that C_1 and C_2 are not symmetric in the game (M, v^C) and the Generalized Solidarity PGI collects this feature assigning different overall power to all coalitions. This is not the case by taking the Generalized Union PGI. As it is the case in this example, both indices typically yield different distributions of power.

Acknowledgments This research received financial support from the Academy of Finland, as well as from Ministerio de Ciencia e Innovación (Spain) and FEDER through Projects ECO2008-03484-C02-02, MTM2011-27731-C03-02 and MTM2011-27731-C03-03 and from Xunta de Galicia through Project INCITE09-207-064-PR.

References

Albizuri, M. J., & Aurrekoetxea, J. (2006). Coalition configurations and the Banzhaf index. *Social Choice and Welfare*, *26*, 571–596.

Albizuri, M. J., Aurrekoetxea, J., & Zarzuelo, J. M. (2006). Configuration values: Extensions of the coalitional Owen value. *Games and Economic Behaviour*, *57*, 1–17.

Alonso-Meijide, J. M., Casas-Méndez, B., Fiestras-Janeiro, M. G., & Holler, M. J. (2010a). Two variations of the public good index for games with a priori unions. *Control and Cybernetics*, *39*, 839–855.

Alonso-Meijide, J. M., Casas-Méndez, B., Fiestras-Janeiro, M. G., Holler, M. J., & Nohn, A. (2010b). Axiomatizations of public good indices with a priori unions. *Social Choice and Welfare*, *35*, 517–533.

Andjiga, N., & Courtin, S. (2010). *Coalition configuration and share functions*. Working paper: CREM, Université de Caen, France.

Banzhaf, J. F. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review*, *19*, 317–343.

Coleman, J. S. (1971), Control of collectivities and the power of a collectivity to act. In: B. Lieberman (Ed.), *Social choice* (pp. 269–300). New York: Gordon and Breach.

Deegan, J., & Packel, E. W. (1978). A new index of power for simple n-person games. *International Journal of Game Theory*, *7*, 113–123.

Holler, M. J. (1982). Forming coalitions and measuring voting power. *Political Studies*, *30*, 262–271.

Holler, M. J., & Nohn, A. (2009). The public good index with threats in a priori unions. *Essays in Honor of Hannu Nurmi, Vol. I, Homo Oeconomicus*, *26*, 393–401.

Holler, M. J., & Packel, E. W. (1983). Power, luck and the right index. *Journal of Economics*, *43*, 21–29.

Johnston, R. J. (1978). On the measurement of power: Some reaction to Laver. *Environment and Planning A*, *10*, 907–914.

Owen, G. (1977). Values of games with a priori unions. In R. Henn & O. Moeschlin (Eds.), *Mathematical economics and game theory* (pp. 76–88). Heidelberg: Springer Verlag.

Shapley, L. S. (1953). A value for n-person games. *Annals of Mathematics Studies*, *28*, 307–318.

Shapley, L. S., & Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review*, *48*, 787–792.

van der Laan, G., & van den Brink, R. (2002). A Banzhaf share function for cooperative games with coalition structure. *Theory and Decision*, *53*, 61–86.

van der Laan, G., & van den Brink, R. (2005). A class of consistent share function for games with coalition structure. *Games and Economic Behavior*, *51*, 193–212.

Winter, E. (1992). The consistency and potential for values with coalition structures. *Games and Economic Behavior*, *4*, 132–144.

Circumstantial Power: Some Hints for Finding Optimal Persuadable or Bribable Voters

Josep Freixas and Montserrat Pons

1 Introduction

Assume that a proposal P has to be submitted to a finite set of voters N, and that each voter i has an independent a priori probability p_i of voting in favor of the proposal.

Suppose now that an external influence might change the perception of some voter i with respect to P, and thus his/her probability of voting for the proposal. In this context, it is of special interest to know which particular voter is more decisive in order to get P approved. The ways to measure whether one voter is more decisive than other can vary, and depend on different factors, but anyone of them must take into account the voting rules and the independent a priori probabilities p_i of the other players with respect to P.

For illustration, let us consider two examples.

Example 1.1 A shareholder company is formed by three majority shareholders a, b_1 and b_2 possessing the 26, 25 and 25 % of the shares, respectively, and, by 24 minority shareholders with 1 % of the shares each. The decisions in the company are taken by absolute majority and each stockholder has as many votes as shares. Assume that the company must vote for an important issue affecting an outsider. This outsider makes an estimation of the voters' preferences and assigns the shareholders to three groups: (A) those who are inclined to vote in favor of his interests; (B) those who are undecided; (C) those who tend to vote against his interests. Suppose that a and b_2 are in group B, b_1 is in A, and all minority

Research partially supported by Grants SGR 2009–1029 of "Generalitat de Catalunya", MTM 2009–08037 from the Spanish Science and Innovation Ministry and MTM 2012–34426 from the Spanish Economy and Competitiveness Ministry.

J. Freixas (✉) · M. Pons
Department of Applied Mathematics III and High Engineering School of Manresa, Technical University of Catalonia, EPSEM, Av. Bases de Manresa, 61-73 E–08242 Manresa, Spain
e-mail: josep.freixas@upc.edu

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*, 315
DOI: 10.1007/978-3-642-35929-3_17,

shareholders are in B except two of them, one in A and the other one in C. The crucial question for the outsider is:

Who are the shareholders that should be addressed the highest persuasion effort?

Example 1.2 Let us suppose that the owners of 20 flats, 4 duplex apartments and 2 shops in a building have to decide if they accept to sell the whole building to a company. In the contract of sale it is established that this kind of decisions must be approved by a majority of 75 % of the owners by taking into account their percentage of the building ownership (each flat has 3 %, each duplex apartment has 6, 5 % and each shop has 7 %). The president of the buyer company has the perception that 10 owners of flats are clearly in favor of accepting the proposal; the owners of 2 duplex apartments and 7 flats are slightly in favor; the owners of 1 shop and 1 duplex apartment are slightly against; and the owners of 1 shop, 1 duplex apartment and 3 flats are clearly against selling their properties.

Assume now that the president of the buyer company is willing to offer extra money to some owner, to ensure his favorable vote and improve in this way his/her expectations of buying the building. Which owner should he select?

In this chapter we present ways to approach the study of the former questions by using two circumstantial power indices: Γ and Ω. This work was initiated in Freixas and Pons (2005), where the problem was solved for the particular case of linear games and under the hypothesis of equal voter's preferences. In a subsequent chapter (Freixas and Pons 2008) we deepened on the persuasion problem and obtained hints on finding optimal persuadable voters for *any* given simple game and for *any* possible distribution of voters' preferences in relation with the proposal at hand. The key idea is the introduction of two additional preorderings defined on the set of voters, which are stronger than the desirability relation, i.e., they imply it, and, as far as we know, had never been considered before in game theory.

In Sect. 2 the Γ and Ω measures are introduced to evaluate the importance of a particular voter in a given context (circumstantial measures). Their main properties are analyzed in Sect. 3. In this section we also formally define the two optimization problems which are considered in the chapter. In Sect. 4 three preorderings are considered and their properties are established. Sections 5 and 6 contain the main results. They are devoted to study the existence of solutions for the optimization problems by using the former preorderings. Section 7 concludes the chapter.

2 Two Measures of Circumstantial Power

2.1 The Framework

A voting system is a simple game $(N, \mathcal{W})$, where $N = \{1, 2, \ldots, n\}$ denotes the set of *players or voters*, subsets of N are *coalitions* and $\mathcal{W}$ is the set of *winning* *coalitions*. Subsets of N that are not in $\mathcal{W}$ are called losing coalitions. A *simple game* is defined to be monotonic: subsets of losing coalitions are again losing. A winning coalition is *minimal* if each proper subset is a losing coalition. The set of minimal winning coalitions is usually denoted by $\mathcal{W}^m$.

Before the votes are cast it is not possible to know which coalition (formed by the "yes" voters) will emerge, but we assume that an estimation p_i is made of the probability that voter i is a member of the supporting coalition, i.e., p_i is the estimated probability of voter i voting in favor of the proposal.

Suppose now that the vote of every voter is *independent* from the vote of the remaining voters. In this case, the probability of the proposal being accepted can be written as

$$f(\mathbf{p}) = \sum_{S \in \mathcal{W}} \prod_{i \in S} p_i \prod_{i \notin S} (1 - p_i). \tag{1}$$

where $\mathbf{p} = (p_1, \ldots, p_n) \in [0, 1]^n$ is said to be the *predictions vector*. The function f is the *multilinear extension* of the simple game $(N, \mathcal{W})$ as introduced by Owen (1972).

Suppose now that an external influence might increase the probability of a player i to vote for the proposal. Assuming that the cost of increasing p_i by Δp_i is proportional to Δp_i but independent of player i and of the value p_i, it will be of interest to know which players are more important in the sense that a change Δp_i on p_i leads to a larger increase of $f(\mathbf{p})$.

It is straightforward to check that the increment on $f(\mathbf{p})$ due to an increment Δp_i on p_i is:

$$\Delta_i f(\mathbf{p}) = f(\mathbf{p} + \Delta_i(\mathbf{p})) - f(\mathbf{p}) = f_i(\mathbf{p}) \Delta p_i \tag{2}$$

where f_i stands for the partial derivative of f with respect to the component i and $\Delta_i(\mathbf{p}) = (0, \ldots, 0, \Delta p_i, 0, \ldots, 0)$.

In this context we will define two measures of the *circumstantial* or *local* power of a voter i, which we will call Γ and Ω, respectively. From our point of view, circumstantial power indicates that such measures do not only depend on the structure of the game but they also depend on the circumstances that encompass each player in each occasion. Such measures are subjective, in the sense that an external observer can consider a certain predictions vector reasonable whereas another considers another one suitable. Thus, each observer is interested in evaluating his/her 'measure' of voting power. Subjectivity of measures has been a topic in literature since Weber (1988) introduced the probabilistic values.

2.2 The Γ Measure of Circumstantial Power

The Γ *power measure* for voter i, in the context $((N, \mathcal{W}), \mathbf{p})$ is defined by:

$$\Gamma_i(\mathbf{p}) = f_i(\mathbf{p}) \tag{3}$$

It is clear that $\Gamma_i(\mathbf{p})$, partial derivative of f with respect to the component i in $\mathbf{p}$, gives a measure of the strategic importance of the voter i in these circumstances. If $\Gamma_i(\mathbf{p})$ is large, small changes in the perception of voter i will give relatively large changes in the probability of P to be approved. Thus, the Γ measure might be appropriate to survey the sensitivity of the game respect to small changes in the voter's perception. Owen (1972; 1975), and Straffin (1977) consider the polynomial expression $f_i(\mathbf{p})$. Napel and Widgrén (2004) give a detailed discussion on measuring sensitivity.

2.3 The Ω Measure of Circumstantial Power

The Ω *power measure* for voter i, in the context $((N, \mathcal{W}), \mathbf{p})$ is defined by:

$$\Omega_i(\mathbf{p}) = (1 - p_i) f_i(\mathbf{p}) \tag{4}$$

Notice that $\Delta_i f(\mathbf{p})$ depends on $f_i(\mathbf{p})$ but also on the values Δp_i. Indeed, it is obvious that if $p_i = 1$ no increase of this probability is possible, while if $p_i = 0$ we can think of an increase $\Delta p_i = 1$. So the potential strategic importance of a player i depends on two factors: the rate of change $f_i(\mathbf{p})$ and the a priori probability p_i. This is why we suggest $\Omega_i(\mathbf{p})$ as a measure of the potential importance of player i.

In some way, $\Omega_i(\mathbf{p})$, expresses the potential increase in the probability of P being accepted due to a change in the perception of voter i. In this sense, if we ask which voters should be bribed in order to get a larger increase in $f(\mathbf{p})$, those voters with the maximum value of $\Omega_i(\mathbf{p})$ would be a good choice. To our knowledge this power voting measure had not been considered before.

2.4 Relationship with Standard Power Indices

Let's recall the Penrose–Banzhaf–Coleman measure and the Shapley–Shubik index, and point out their relationship with Γ and Ω. The *Penrose–Banzhaf–Coleman measure of voting power* (see Banzhaf 1965; Coleman 1971; Penrose 1946) is

$$\psi_i = 2^{1-n} \eta_i(N, \mathcal{W})$$

where $\eta_i(N, \mathcal{W})$ stands for the number of coalitions in which i is crucial, i.e.

$$\eta_i(N, \mathcal{W}) = | \{ S \mid i \in S, S \in \mathcal{W}, S \setminus \{i\} \notin \mathcal{W} \mid .$$

The *Shapley–Shubik power index* (see Dubey (1975), Shapley (1953; 1962) and Shapley and Shubik (1954)) is

$$\phi_i = \sum_{\{S \mid i \text{ iscrucialin S}\}} \frac{(s-1)!(n-s)!}{n!}$$

where $|S| = s$ and $|N| = n$. For more material on both indices see, for example, Felsenthal and Machover (1998).

Both indices can be written in terms of the partial derivative f_i of f:

$$(a)\ \phi_i = \int_0^1 f_i(p, p, \ldots, p)dp$$

$$(b)\ \psi_i = f_i(\frac{1}{2}, \frac{1}{2}, \ldots, \frac{1}{2})$$

for all $i \in N$ (See Owen (1972, 1975), respectively). A nice generalization has been given by Straffin (1988): by considering multiple integration on the unit cube $[0, 1]^n$, it follows from Fubini's theorem that, for all i,

$$\psi_i = \int_{[0,1]^n} f_i(p_1, p_2, \ldots, p_n)dp_1\, dp_2 \ldots, dp_n.$$

We see that the two measures are closely related to the circumstantial power measure Γ. With respect to the Ω measure, it can be proved that

$$\phi_i = \frac{\int_0^1 \Omega_i(p, p, \ldots, p)dp}{\int_0^1 f(p, p, \ldots, p)dp}$$

3 Some Properties of Γ and Ω

In this section we will point out some properties of Γ and Ω that show their behavior as power measures. We already discussed these properties in Freixas and Pons (2005). Some of them depend on properties of the multilinear function f defined in (1), so we start by stating its monotonicity. This is an important feature because it implies that any improvement of the perception of any single voter (with the other components remaining unchanged) produces an increase in the game's expectation for the proposal to be accepted.

Property 3.1 *The multilinear function f is non-decreasing in each variable.*

Our concern now is to compare the measures Γ and Ω for two arbitrary voters. To this end we will use the so-called factoring algorithm (pivot decomposition method). The basic idea of this method is to make a conditional probability argument using the relationship

$$f(\mathbf{p}) = p_i f(1_i, \mathbf{p}) + (1 - p_i) f(0_i, \mathbf{p}) \tag{5}$$

where the expression $(x_i, \mathbf{p})$ stands for $(p_1, \ldots, p_{i-1}, x, p_{i+1}, \ldots, p_n)$ if $\mathbf{p} = (p_1, \ldots, p_n)$ and $1 \leq i \leq n$. Formula (5) follows from the law of total probability.

Lemma 3.2 *For all $i \in N$ it is*

(i) $\Gamma_i(\mathbf{p}) = f(1_i, \mathbf{p}) - f(0_i, \mathbf{p})$,
(ii) $\Omega_i(\mathbf{p}) = f(1_i, \mathbf{p}) - f(\mathbf{p})$.

Corollary 3.3

(i) $0 \leq \Gamma_i(\mathbf{p}) \leq 1$ *for all* $i \in N$,
(ii) $0 \leq \Omega_i(\mathbf{p}) \leq 1$ *for all* $i \in N$.

Notice that $\Gamma_i(\mathbf{p}) = 1$ iff $\mathcal{W}^m = \{\{i\}\}$, and $\Omega_i(\mathbf{p}) = 1$ iff $\mathcal{W}^m = \{\{i\}\}$ and $\mathbf{p} = \mathbf{0}$.

Proposition 3.4 *For $i, j \in N$ it holds*

$$\begin{aligned}
\Gamma_i(\mathbf{p}) - \Gamma_j(\mathbf{p}) &= (p_j - p_i)[f(1_i, 1_j, \mathbf{p}) + f(0_i, 0_j, \mathbf{p}) - f(1_i, 0_j, \mathbf{p}) - f(0_i, 1_j, \mathbf{p})] \\
&\quad + [f(1_i, 0_j, \mathbf{p}) - f(0_i, 1_j, \mathbf{p})], \\
\Omega_i(\mathbf{p}) - \Omega_j(\mathbf{p}) &= p_j[f(1_i, 1_j, \mathbf{p}) - f(1_i, 0_j, \mathbf{p})] - p_i[f(1_i, 1_j, \mathbf{p}) - f(0_i, 1_j, \mathbf{p})] \\
&\quad + [f(1_i, 0_j, \mathbf{p}) - f(0_i, 1_j, \mathbf{p})].
\end{aligned}$$

Proposition 3.4 will be a useful tool to analyze the optimization problems motivated in the introduction. We remark that the outsider, in this first take on the general problem, is restricted to approach only one voter rather than dividing a fixed amount of persuasion effort or bribe between several ones.

Let us formally introduce the optimization problems:

Persuasion problem:
Given $(N, \mathcal{W})$ and $\mathbf{p} \in [0,1]^n$, let $\epsilon > 0$ be small enough and $\epsilon_i = (0, \ldots, 0, \epsilon, 0, \ldots, 0)$ for $i = 1, \ldots, n$. Consider the set $S(\mathbf{p}) = \{\mathbf{p} + \epsilon_i,\ i = 1, \ldots, n\}$. Notice that $S(\mathbf{p})$ is formed by those vectors obtained from $\mathbf{p}$ by slightly increasing one of its components and leaving the rest unchanged.

Problem 1 is to find a voter i such that

$$f(\mathbf{p} + \epsilon_i) = \max_{\mathbf{r} \in S(\mathbf{p})} f(\mathbf{r}).$$

Bribe problem:
Given $(N, \mathcal{W})$ and $\mathbf{p} \in [0,1)^n$, let $\mathbf{m}_i = (0, \ldots, 0, 1 - p_i, \ldots, 0)$ for $i = 1, \ldots, n$. Consider the set $B(\mathbf{p}) = \{\mathbf{p} + \mathbf{m}_i, i = 1, \ldots, n\}$. Notice that $B(\mathbf{p})$ is formed by those vectors obtained from $\mathbf{p}$ by replacing one of its components by 1 and leaving the rest unchanged.

Problem 2 is to find a voter i such that

$$f(\mathbf{p} + \mathbf{m}_i) = \max_{\mathbf{r} \in B(\mathbf{p})} f(\mathbf{r}).$$

These two problems have an equivalent form to be stated using indices Γ and Ω.

Proposition 3.5

(i) *A voter i is a solution for the persuasion problem if and only if*

$$\Gamma_i(\mathbf{p}) = \max_{1 \le j \le n} \Gamma_j(\mathbf{p}).$$

(ii) *A voter i is a solution for the bribe problem if and only if*

$$\Omega_i(\mathbf{p}) = \max_{1 \le j \le n} \Omega_j(\mathbf{p}).$$

The answer to these optimization problems depends, of course, on the particular game $(N, \mathcal{W})$ and on the predictions vector $\mathbf{p}$.

It is clear that if the number of voters is small and a predictions vector $\mathbf{p}$ is fixed, the easiest way of solving the considered optimization problems would consist of finding the n partial derivatives of f, evaluating the desired measure in $\mathbf{p}$, and choosing the voter for which this value is maximum. But in real—world problems it is usually more meaningful to estimate a ranking of voter's preferences rather than to assign exact values to the components of $\mathbf{p}$. In such a case it is obviously impossible to calculate the exact value of f and its partial derivatives.

This chapter aims to analyze the problem for *any* simple game $(N, \mathcal{W})$ and for *any* predictions vector $\mathbf{p}$, by reviewing the results given in Freixas and Pons (2005; 2008). The proposed approach gives an answer to the problem whenever the set of minimal winning coalitions and the ranking between the components of $\mathbf{p}$ are known. It is possible to solve these optimization problems by implementing an algorithm in a computer. From now on, $(N, \mathcal{W})$ is assumed to be a simple game, $\mathcal{W}^m$ its set of minimal winning coalitions, f its multilinear extension and $\mathbf{p} \in (0,1)^n$ a predictions vector.

4 Some Useful Preorderings on N

Definition 4.1 We use the set of winning coalitions $\mathcal{W}$ to define different binary relations between two elements $i, j \in N$

(i) The **external subordination relation**. $i \models j$ if and only if $i = j$or$[S \in \mathcal{W}, j \in S, i \notin S \Rightarrow S \setminus \{j\} \in \mathcal{W}]$ If $i \models j$ we say that voter i *externally subordinates* j.
(ii) The **internal subordination relation**. $i \trianglerighteq j$ if and only if $i = j$or $[S \in \mathcal{W}, i, j \in S \Rightarrow S \setminus \{j\} \in \mathcal{W}]$ If $i \trianglerighteq j$ we say that voter i *internally subordinates* j.
(iii) The **desirability relation**. $i \succeq j$ if and only if $[S \cup \{j\} \in \mathcal{W} \Rightarrow S \cup \{i\} \in \mathcal{W}$, whenever $S \subseteq N \setminus \{i, j\}]$ If $i \succeq j$ we say that voter i is *at least as desirable as* j.

If i externally subordinates j then for any winning coalition S which does not contain i, the fact of S containing j or not does not change S' condition of winning. Similarly, if i internally subordinates j then for any winning coalition S which contains i, the fact of S containing j or not does not change S' condition of winning. Finally, if i is at least as desirable as j then i can be put instead of j in any winning coalition S without changing S' condition of winning.

As far as we know the two subordination relations introduced in Freixas and Pons (2008) are new in the context of voting systems. However, the desirability relation goes back at least to Isbell (1956). It is clear that all these relations are reflexive. The following proposition states their transitivity, which is a known property in the case of the desirability relation. The proofs of all the properties in this section can be found in Freixas and Pons (2008).

Proposition 4.2 *The external subordination relation, the internal subordination relation and the desirability relation, are preorderings on* N.

The following proposition shows that the former preorderings can be expressed in terms of minimal winning coalitions.

Proposition 4.3 *Let* i, j *be different elements in* N.

(i) $i \models j$ *if and only if* $[S \in \mathcal{W}^m, j \in S \Rightarrow i \in S]$
(ii) $i \trianglerighteq j$ *if and only if* $[S \in \mathcal{W}^m, j \in S \Rightarrow S \cup \{i\} \setminus \{j\} \in \mathcal{W}]$
(iii) $i \succeq j$ if and only if $[S \in \mathcal{W}^m, j \in S, i \notin S \Rightarrow S \cup \{i\} \setminus \{j\} \in \mathcal{W}]$

From this proposition it is clear that if i, j are different elements in N and $(N, \mathcal{W})$ is a game such that $\{i\} \in \mathcal{W}^m$ then: $i \triangleright j, i \succeq j$ and $j \not\models i$. If, moreover, there is in $\mathcal{W}^m$ some other minimal winning coalition containing j, then $i \not\models j$. However, if $\mathcal{W}^m = \{\{i\}\}$ i.e., $(N, \mathcal{W})$ is the dictatorship of voter i, then: $i \triangleright j$, $i \succ j$ and $i \vdash j$ for any other voter j.

It is also easy to see that if $(N, \mathcal{W})$ is a game such that there exists $S \in \mathcal{W}^m$ with $i, j \in S$ then $i \not\trianglerighteq j$ and $j \not\trianglerighteq i$.

Each one of the above preorderings induces an equivalence relation on the set of voters N, and an order relation on the corresponding set of equivalence classes. Let us introduce some notation.

Definition 4.4 Given two elements $i, j \in N$, the following binary relations will be used:

(1.a) $i \vdash j$ if and only if $i \vDash j$ and $j \nvDash i$

(1.b) $i \vDash\!\dashv j$ if and only if $i \vDash j$ and $j \vDash i$
(2.a) $i \triangleright j$ if and only if $i \trianglerighteq j$ and $j \not\trianglerighteq i$
(2.b) $i \bowtie j$ if and only if $i \trianglerighteq j$ and $j \trianglerighteq i$
(3.a) $i \succ j$ if and only if $i \succeq j$ but $j \not\succeq i$
(3.b) $i \equiv j$ if and only if $i \succeq j$ and $j \succeq i$

The binary relation $\vdash$ is closely related to the concept of inferior player introduced by Napel and Widgrén (2001). Specifically, it is not difficult to prove that a player i is inferior in a game $(N, \mathcal{W})$ (as defined in Napel and Widgrén (2001)) if and only if there exist $j \in N$ $(j \neq i)$ such that $j \vdash i$.

These binary relations are not independent. The following Proposition 4.5 states some relations among them. Notice that part (v) is a characterization for null voters in terms of the two subordination relations.

Proposition 4.5 *Let i, j be different elements in N. Then,*

(i) $i \vDash j \quad \Rightarrow \quad i \succeq j$
(ii) $i \vDash j \quad \Rightarrow \quad i \succ j$
(iii) $i \trianglerighteq j \quad \Rightarrow \quad i \succeq j$
(iv) $i \trianglerighteq j \quad \Rightarrow \quad i \succ j$
(v) $i \vDash j$ *and* $i \trianglerighteq j \Leftrightarrow j$ *is a null voter*

The following Proposition 4.6 states that if i votes "no" and externally subordinates j, then the probability of the proposal being accepted does not depend on the vote of j. The same happens if i votes "yes" and internally subordinates j. On the other hand, if i is at least as desirable as j, then the conditional probability of the proposal being accepted assuming that i votes "yes" and j votes "no" is greater than the probability of acceptance, if their votes are interchanged.

Proposition 4.6 *Given two different elements $i, j \in N$,*

(i) $i \vDash j \quad \Leftrightarrow \quad f(0_i, 0_j, \mathbf{p}) = f(0_i, 1_j, \mathbf{p})$.
(ii) $i \trianglerighteq j \quad \Leftrightarrow \quad f(1_i, 0_j, \mathbf{p}) = f(1_i, 1_j, \mathbf{p})$.
(iii) $i \succeq j \quad \Rightarrow \quad f(1_i, 0_j, \mathbf{p}) - f(0_i, 1_j, \mathbf{p}) \geq 0$.
(iv) $i \succ j \quad \Rightarrow \quad f(1_i, 0_j, \mathbf{p}) - f(0_i, 1_j, \mathbf{p}) > 0$.

5 Using the Desirability Relation

The results in this section show how the desirability relation can be used for finding optimal persuadable and optimal bribable voters. The corresponding proofs can be found in Freixas and Pons (2005). We aim at selecting either optimal persuadable or optimal bribable voters in a voting context $((N, \mathcal{W}), \mathbf{p})$. We assume that the only information we have on the predictions vector $\mathbf{p}$ is the ranking between its components. Since null voters have both Γ and Ω measures zero for any predictions vector $\mathbf{p}$, they will always be discarded as optimal voters. This is why, from now on, we assume that all voters are not null.

Let us start with the bribe problem. The following Theorem 5.1 states that if the prediction for one voter is smaller (or equal) than the prediction for another one, who is less desirable than him, then the first one is a better candidate to be bribed, and the second one can be discarded as optimal bribable voter. Furthermore, if the prediction for one voter is strictly smaller than the prediction for another one, who is more desirable than him, then none of them can be discarded as a candidate to be bribed.

Theorem 5.1 *Let i, j be different elements in N. Then, the Ω measure satisfies:*

(i) $i \succeq j,\ p_i \leq p_j \Rightarrow \Omega_i(\mathbf{p}) \geq \Omega_j(\mathbf{p})$,
(ii) $i \succ j,\ p_i \leq p_j \Rightarrow \Omega_i(\mathbf{p}) > \Omega_j(\mathbf{p})$,
(iii) $i \succeq j,\ p_i < p_j \Rightarrow \Omega_i(\mathbf{p}) > \Omega_j(\mathbf{p})$,
(iv) *if $i \succeq j$ and $p_i > p_j$, then it is possible to find a predictions vector $\mathbf{p}^*$ such that $p_i{}^* > p_j{}^*$ and the sign of $\Omega_i(\mathbf{p}) - \Omega_j(\mathbf{p})$ differs from the sign of $\Omega_i(\mathbf{p}^*) - \Omega_j(\mathbf{p}^*)$*

As a consequence of Theorem 5.1, the optimal bribable voter is the maximum one with respect to the desirability relation if the game is linear, i.e., the desirability relation is total, and the ranking of predictions reverses the ranking given by the desirability relation.

Corollary 5.2 *Let $(N, \mathcal{W})$ be a linear simple game with $1 \succeq 2 \succeq \ldots \succeq n$, and assume that $p_1 \leq p_2 \leq \ldots \leq p_n$. Then the Ω measure satisfies*

$$\Omega_1(\mathbf{p}) \geq \Omega_2(\mathbf{p}) \geq \ldots \geq \Omega_n(\mathbf{p}).$$

Any inequality $\Omega_i(\mathbf{p}) \geq \Omega_{i+1}(\mathbf{p})$ can be replaced by $\Omega_i(\mathbf{p}) > \Omega_{i+1}(\mathbf{p})$ if either $p_i < p_{i+1}$ or $i \succ i+1$.

The following Theorem 5.3 states that if two voters have *the same* prediction then the more desirable one is a better candidate to be persuaded. (The case of two voters having different predictions will be analyzed in the next section.)

Theorem 5.3 *Let i, j be different elements in N. Then the Γ measure satisfies*:

(i) $i \succeq j$ *and* $p_i = p_j \quad \Rightarrow \quad \Gamma_i(\mathbf{p}) \geq \Gamma_j(\mathbf{p})$
(ii) $i \succ j$ *and* $p_i = p_j \quad \Rightarrow \quad \Gamma_i(\mathbf{p}) > \Gamma_j(\mathbf{p})$

From Theorems 5.1 and 5.3 it is obvious that if two voters are related by the desirability order and the predictions for both voters are the same, then the voter with a highest potential to improve the game's expectation to pass the proposal is (for both optimization problems) the most powerful voter with respect to the desirability relation. For linear games and predictions vectors with equal components we may exploit this property to the full extent.

Theorem 5.4 *Let $(N, \mathcal{W})$ be a linear simple game with $1 \succeq 2 \succeq \ldots \succeq n$, and $p_k = p \in (0,1)$ for all $k = 1, \ldots, n$. Then*

(i) $\Gamma_1(\mathbf{p}) \geq \Gamma_2(\mathbf{p}) \geq \ldots \geq \Gamma_n(\mathbf{p})$.
(ii) $\Omega_1(\mathbf{p}) \geq \Omega_2(\mathbf{p}) \geq \ldots \geq \Omega_n(\mathbf{p})$.

In both cases the inequality '$\geq$' can be replaced by '$=$' if $i+1 \equiv i$, and by '$>$' if $i \succ i+1$.

6 Deepening on the Persuasion Problem

In this section we use the external and the internal subordination preorderings for comparing the Γ-measure of two voters with different predictions. Theorem 6.1 states that in certain situations there are no general arguments to discard any voter as a candidate to be persuaded. Specifically, three different relations between two voters are identified under which anyone of them has the greatest Γ-measure if a convenient predictions vector is chosen. The proofs of all theorems in this section can be found in Freixas and Pons (2008).

Theorem 6.1 *Let i, j be different elements in N, and $x, y \in [0,1]$. Then,*

(i) *If $i \unrhd j$ and $i \nvDash j$, there exists a predictions vector $\mathbf{p} \in [0,1]^n$ such that $p_i = x$, $p_j = y$ and $\Gamma_i(\mathbf{p}) - \Gamma_j(\mathbf{p}) = x - y$.*
(ii) *If $i \unrhd j$ and $i \ntrianglerighteq j$, there exists a predictions vector $\mathbf{p} \in [0,1]^n$ such that $p_i = x$, $p_j = y$ and $\Gamma_i(\mathbf{p}) - \Gamma_j(\mathbf{p}) = y - x$.*
(iii) *If $i \not\unrhd j$, and $x, y \in [0,1]$, there exists a predictions vector $\mathbf{p} \in [0,1]^n$ such that $p_i = x$, $p_j = y$ and $\Gamma_j(\mathbf{p}) - \Gamma_i(\mathbf{p}) = 1$.*

The following Theorem 6.2 states that if a voter i externally subordinates another voter j, and its prediction p_i is smaller than p_j, then voter i is a better candidate than j to be persuaded. But if p_i is greater than p_j then none of them can be discarded as a candidate to be persuaded.

Theorem 6.2 *Let i, j be different elements in N. Then,*

(i) *$i \vDash j$ and $p_i \leq p_j \quad \Rightarrow \quad \Gamma_i(\mathbf{p}) \geq \Gamma_j(\mathbf{p})$*
(ii) *$i \vDash j$ and $p_i < p_j \quad \Rightarrow \quad \Gamma_i(\mathbf{p}) > \Gamma_j(\mathbf{p})$*
(iii) *If $i \vDash j$ and $p_i > p_j$, there exist vectors $\mathbf{p}^*, \mathbf{p}^\diamond \in [0,1]^n$ with $p_i^* = p_i^\diamond = p_i$ and $p_j^* = p_j^\diamond = p_j$, such that $\Gamma_i(\mathbf{p}^*) < \Gamma_j(\mathbf{p}^*)$ and $\Gamma_i(\mathbf{p}^\diamond) > \Gamma_j(\mathbf{p}^\diamond)$*

Theorem 6.3 states that if a voter i internally subordinates another voter j, and its prediction p_i is greater than p_j, then voter i is better candidate than j to be persuaded. However, if p_i is smaller than p_j then none of them can be discarded as a candidate to be persuaded.

Theorem 6.3 *Let i, j be different elements in N. Then,*

(i) $i \trianglerighteq j$ *and* $p_i \geq p_j \quad \Rightarrow \quad \Gamma_i(\mathbf{p}) \geq \Gamma_j(\mathbf{p})$
(ii) $i \trianglerighteq j$ *and* $p_i > p_j \quad \Rightarrow \quad \Gamma_i(\mathbf{p}) > \Gamma_j(\mathbf{p})$
(iii) *If* $i \triangleright j$ *and* $p_i < p_j$, *there exist vectors* $\mathbf{p}^*, \mathbf{p}^\diamond \in [0,1]^n$ *with* $p_i^* = p_i^\diamond = p_i$ *and* $p_j^* = p_j^\diamond = p_j$, *such that* $\Gamma_i(\mathbf{p}^*) < \Gamma_j(\mathbf{p}^*)$ *and* $\Gamma_i(\mathbf{p}^\diamond) > \Gamma_j(\mathbf{p}^\diamond)$

Theorem 6.4 states that if a voter i is at least as desirable as another voter j, and prediction p_i is different than p_j then none of them can be discarded as a candidate to be persuaded, except in the cases provided by the above theorems.

Theorem 6.4 *Let i, j be different elements in N. Then,*

If $i \succ j, i \not\models j, i \not\triangleright j$ *and* $p_i \neq p_j$, *there exist vectors* $\mathbf{p}^*, \mathbf{p}^\diamond \in [0,1]^n$ *with* $p_i^* = p_i^\diamond = p_i$ *and* $p_j^* = p_j^\diamond = p_j$, *such that* $\Gamma_i(\mathbf{p}^*) < \Gamma_j(\mathbf{p}^*)$ *and* $\Gamma_i(\mathbf{p}^\diamond) > \Gamma_j(\mathbf{p}^\diamond)$

7 Some Final Comments

In order to illustrate how the theorems can be used to get the desired solutions, we revisit the introductory examples in the light of the results provided in this chapter.

Example 7.1 (Example 1.1 revisited)

Let a be the main shareholder with 26 % of the shares, b_1 and b_2 those shareholders that have a 25 % each, and $\{c_i : 1 \leq i \leq 24\}$ the set of minority shareholders with 1 % of the shares each one of them. An outsider estimates that in relation to his interests the shareholders are positioned as follows:

(A) Quite in favor of his interests (with prediction p_A, being $p_A < 1$): b_1 and c_1.
(B) Undecided voters (with prediction p_B near $1/2$, being $p_B < p_A$): a, b_2, and $\{c_i : 2 \leq i \leq 23\}$.
(C) Quite against his interests (with prediction p_C, being $0 < p_C < p_B$): c_{24}.

From the established voting rules, it is easy to check that the minimal winning coalitions are: $\{a, b_1\}$, $\{a, b_2\}$ and $\{b_1, b_2, c_k\}$ for any k $(1 \leq k \leq 24)$. From proposition 4.3 we deduce that

$$a \triangleright c_k \qquad \text{for each } k = 1, \ldots, 24. \tag{6}$$

$$a \succ b_j \qquad \text{for each } j = 1, 2. \tag{7}$$

$$b_j \models c_k \qquad \text{for each } j = 1, 2 \ \text{ and } k = 1, \ldots, 24. \tag{8}$$

From (6) and Theorem 6.3, c_{24} should be discarded as a candidate to be persuaded. Also from (6), and using Proposition 4.5, it is $a \succ c_k$ for $k = 1, \ldots, 24$ and therefore, by Theorem 6.4, all other c_j, except c_1, must also be discarded. But, taking into account (8), Theorem 6.2 let us discard also c_1 as optimal persuadable voter. Finally, from (7) and Theorem 6.4, b_2 should also be discarded as a candidate to be persuaded. In conclusion, none of the 24 minority shareholders must be taken into account as candidates to be persuaded, and, although voters b_1 and b_2 possess the same number of shares, the first one is more crucial in the context described than the second one. In conclusion, the set of optimal persuadable voters is in this case: $\{a, b_1\}$. We cannot decide which one of them is the preferred target if the particular values of their respective predictions, p_B and p_A are not known. □

Example 7.2 (Example 1.2 revisited)

Let a_1, a_2, a_3, a_4 be the owners of duplex apartments, b_1, b_2 the owners of shops and $\{c_i : 1 \le i \le 20\}$ the owners of flats. The president of the buyer company estimates the opinion of each owner in the following way:

(A) Clearly in favor of selling (with prediction $p_A > 0.9$): $\{c_i : 11 \le i \le 20\}$.
(B) Slightly in favor of selling (with prediction p_B with $0.5 < p_B < 0.9$): a_3, a_4, and $\{c_i : 4 \le i \le 10\}$.
(C) Slightly against selling (with prediction p_C, being $0.2 < p_C < 0.5$): b_2 and a_2
(D) Clearly against selling (with prediction p_D, being $p_D < 0.2$): b_1, a_1, c_1, c_2, c_3.

Since the buyer wants to ensure the favorable vote of one of the owners, he should select the one with maximum Ω measure. Theorem 5.1 allows us to deduce who is this owner. In this case the desirability relation is total: $b_1 \equiv b_2$, $a_1 \equiv a_2 \equiv a_3 \equiv a_4$, $c_i \equiv c_j$ for $1 \le i, j \le 20$, and $b_k \succ a_m \succ c_i$ for $k = 1, 2; m = 1, 2, 3, 4; 1 \le i \le 20$. The fact that $\{c_i : 1 \le i \le 3\}$ are equivalent by the desirability relation and have the same prediction (p_D) implies that they have the same Ω measure. As a consequence, we will only refer to c_1 as a representative of these three owners. In an analogous way, c_4 will represent $\{c_i : 4 \le i \le 10\}$ (all of them with prediction p_B) and c_{11} will represent $\{c_i : 11 \le i \le 20\}$ (all of them with prediction p_A). From Theorem 5.1 we have:

$$
\begin{array}{lll}
b_1 \succeq b_2, & p_D < p_C & \Rightarrow \Omega_{b_1} > \Omega_{b_2} \\
a_1 \succeq a_2 \succeq a_3 \succeq a_4, & p_D < p_C < p_B & \Rightarrow \Omega_{a_1} > \Omega_{a_2} > \Omega_{a_3} = \Omega_{a_4} \\
c_1 \succeq c_4 \succeq c_{11}, & p_D < p_B < p_A & \Rightarrow \Omega_{c_1} > \Omega_{c_4} > \Omega_{c_{11}} \\
b_1 \succ a_1, & \text{and their predictions are equal} & \Rightarrow \Omega_{b_1} > \Omega_{a_1} \\
b_1 \succ c_1, & \text{and their predictions are equal} & \Rightarrow \Omega_{b_1} > \Omega_{c_1}
\end{array}
$$

Thus, the owner who has maximum Ω measure is b_1. He should be selected to be offered more money to ensure his favorable vote.

8 Conclusion

The chapter proposes new instruments for measuring power of individuals in binary decision schemes. Contrarily to the traditional analysis, we do not consider a static model in which the voting rules and the voters' acceptance probabilities are fixed. Instead, we assume that the outsider's estimation of voters' probabilities can change. In this context, we intend to answer the following important question:

'Under the assumption that the effort to convince an individual voter to be in favor of a proposal is always the same, and each voter decides independently of the others, which one is the "best" to be persuaded or bribed in order to maximize the probability for the proposal to be approved?'

In order to help answering this question, two power measures are considered. The first one is Straffin's power polynomial, denoted by Γ. The Γ measure is appropriate to survey the sensitivity of the game respect to small changes in the voter's perception. Voters for which Γ is a maximum are those who cause a larger change in the probability to pass the proposal at hand when they slightly change their individual perceptions.

The second measure, Ω, tries to evaluate the potential strategic importance of each voter in order to make the desired output more likely to be achieved. This measure multiplies Γ by a term capturing the degree to which an outsider interested in passage of a proposal could still increase a given voter's acceptance probability. Voters for which Ω is maximum are those that produce a greater change in the probability to pass the proposal when they ensure their vote for it. Bribes might be offered to these voters.

The main results on the proposed optimization problems are obtained by using three preorderings on the set of voters. The theorems in this chapter provide us with sufficient conditions for comparing the Γ measure or the Ω measure of two given voters. More precisely, the theorems state that if two voters i, j are comparable by one of these pre-orderings, and if their predictions p_i, p_j satisfy an adequate (in)equality, then we can determine a priori the ordering between their measures without the need of computing them. As a consequence, we can use these theorems to select a list of voters to be persuaded or bribed, given any particular ranking of their predictions. Let's note that this procedure can be easily implemented in a computer.

References

Banzhaf, J. F. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review*, *19*, 317–343.

Coleman, J. S. (1971). Control of collectivities and the power of a collectivity to act. In B. Lieberman (Ed.), *Social choice* (pp. 269–300). New York: Gordon and Breach.

Dubey, P. (1975). On the uniqueness of the Shapley value. *American Political Science Review*, *4*, 131–139.

Felsenthal, D. S., & Machover, M. (1998). *The measurement of voting power: Theory and practice, problems and paradoxes*. Cheltenham: Edward Elgar.

Freixas, J., & Pons, M. (2005). Two measures of circumstantial power: Influences and bribes. *Homo Oeconomicus*, *22*, 569–588.

Freixas, J., & Pons, M. (2008). Circumstantial power: Optimal persuadable voters. *European Journal of Operational Research*, *186*, 1114–1126.

Isbell, J. R. (1956). A class of majority games. *Quarterly Journal of Mathematics Oxford Series*, *7*(2), 183–187.

Napel, S., & Widgrén, M. (2001). Inferior players in simple games. *International Journal of Game Theory*, *30*, 209–220.

Napel, S., & Widgrén, M. (2004). Power measurement as sensitivity analysis: A unified approach. *Journal of Theoretical Politics*, *16*, 517–538.

Owen, G. (1972). Multilinear extensions of games. *Management Science*, *18*, 64–79.

Owen, G. (1975). Multilinear extensions and the Banzhaf value. *Naval Research Logistics Quarterly*, *22*, 741–750.

Penrose, L. S. (1946). The elementary statistics of majority voting. *Journal of the Royal Statistical Society*, *109*, 53–57.

Shapley, L. S. (1953). A value for n-person games. In A. W. Tucker & H. W. Kuhn (Eds.), *Contributions to the theory of Games II* (pp. 307–317). Princeton: Princeton University Press.

Shapley, L. S. (1962). Simple games: An outline of the descriptive theory. *Behavioral Science*, *7*, 59–66.

Shapley, L. S., & Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review*, *48*, 787–792.

Straffin, P. D. (1977). Homogeneity, independence and power indices. *Public Choice*, *30*, 107–118.

Straffin, P. D. (1988). The Shapley-Shubik and Banzhaf power indices. In Roth, A. E. (Ed.), *The Shapley value: Essays in honor of Lloyd S. Shapley* (pp 71–81). Cambridge: Cambridge University Press.

Weber, R. J. (1988). Probabilistic values for games. In A. W. Tucker & H. W. Kuhn (Eds.), *Contributions to the theory of Games II* (pp. 101–119). Princeton: Princeton University Press.

Part IV
Applications of Voting Power Measures

Power Indices and the Design of Electoral/ Constitutional Systems

Ron Johnston

1 Introduction

There are three main types of social science research. First, there is that which is oriented at a closed, usually small, group of scholars, which may influence both the nature of their personal academic development and that of their academic discipline, but little else—the results are almost all published in academic journals. Secondly, there is research that is 'applied' in nature, by which is usually meant that it is undertaken, perhaps under contract to a sponsor, to influence programmes for change in the aspect of the world being studied—the results are mainly confined to consultancy reports, many of which are treated as confidential and circulated to restricted audiences only. Finally, there is research which is emancipatory in its goals, produced by academics who hope, through their educational and other activities, to influence how people appreciate the world, and so plan to change it—the results are published in professional, trade and other journals, aimed at a readership outside academia. Those three categories, which classify researchers as scholars, technocrats, and emancipators respectively, are necessarily coarse, and much of their work overlaps two, if not all three. Nevertheless, they provide a structure for this largely polemical piece on the perceived overly-academic orientation of work on power indices and related issues.

There is a substantial research literature on the measurement of power, applying a wide range of indices to studies of coalition-forming and similar behaviour. Most of it falls into the first of the three categories above: I have little evidence of much that can be characterized as 'applied research'—which some may consider a 'good

An earlier version of this paper has been published in *Power Measures. Volume I* (Homo Oeconomicus 17), edited by Manfred J. Holler and Guillermo Owen (2000).

R. Johnston (✉)
School of Geographical Sciences, University of Bristol, University Road,
Bristol, BS8 ISS, UK
e-mail: r.johnston@bristol.ac.uk

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*, 333
DOI: 10.1007/978-3-642-35929-3_18, © Springer-Verlag Berlin Heidelberg 2013

thing' because they believe that much of the work on power indices is fundamentally flawed (see, for example, Morris 1996).[1] Nor has there been much in the third category, perhaps because most of the research literature uses a sophisticated technical language that the great majority of social science undergraduates find impenetrable (and probably do not wish to penetrate, dismissing it as overly-technical and unrepresentative of the real world they live in and want to change).[2] Finally, there is little evidence of work being published outside the in-literature of academic disciplines, in journals and other fora whose goal is to influence popular opinion.[3] We have been active as scholars, but much less so as technocrats and emancipators.

There is a crying need for general education about the issues concerning the design and operation of electoral and governmental systems addressed by the power indices literature, for writing that stimulates increased awareness about the issues explored in such depth by quantitative analysts of power and its exercise. This is illustrated by three examples from my own recent experience.

2 Electoral Reform in New Zealand

In 1992 and 1993 the New Zealand electorate voted to replace the country's first-past-the-post (fptp) electoral system for its unicameral Parliament by one based on the German hybrid combination of fptp and multi-member electorates using the list system (mixed member proportional—MMP). This shift was proposed by a Royal Commission report on the country's electoral system (Royal Commission 1986); it came about because of general voter dissatisfaction with the policies of both of the

[1] It could be argued that because no system so far designed by academics is perfect then none should be advocated. However, if the current system is flawed (which must be the case if no system is perfect) then decision-makers and opinion-leaders should be made aware of alternatives to it, which could remove some of its flaws even though they may introduce others (which may be more acceptable than those currently experienced). If academics remain in 'ivory towers' they can hardly complain if democracy continues to fail on key criteria or if new systems are introduced which are fundamentally flawed. Universities were presented in a recent review of higher education in the UK (*The Dearing Report*) as, inter alia, 'the conscience of the nation'—but they cannot fulfil that role if they isolate themselves from the 'messiness' and imperfections of the 'real world'.

[2] This is certainly the case with the social science disciplines I know best—geography, political science and sociology—where there has been a substantial reaction against quantitative work in recent years (see Bechofer 1996, but also Riba 1996), perhaps especially so in the United Kingdom. Economics and psychology are much less affected by this trend, but I have no evidence that their practitioners have any major 'applied' influence in the field being discussed here.

[3] An excellent example of the small body of work which is both scholarly and written to persuade a wider audience is Steven Brams' advocacy of approval voting (Brams and Fishburn 1983).

country's main political parties in the 1980s and early 1990s and was expressed, against the advice of those parties, in two referendums on the possibility of a change offered by the then Prime Minister.[4] There was considerable debate over the desirability of change during the run-up to the referendums (especially the second, binding, referendum in 1993), accompanied by a government-sponsored public education programme (see Catt et al. 1992, for example), but little if any discussion of the important, and well-established within academia,[5] finding of most analyses of proportional representation (*PR*): that *PR* does not usually lead to proportional power (*PP*).

The country's political system was in considerable chaos in the two years prior to the first election under its new multi-member proportional (MMP) system, as parties split and new alliances were formed in the search for political influence. Boston et al. (1996) note that the country successively experienced single-party majority government, coalition majority government, coalition minority government, and single-party minority government over a three-year period, during which there was no general election. The party system within the House of Representatives fragmented as groups split from their parents in order to mobilize separate electoral support and, hopefully, enhance their bargaining power after the first MMP election: four parties were elected to the House in 1993; by 1995, without any election being held, it had seven.[6]

The first general election under the new system was held in mid-October 1996, with no party obtaining a majority of the seats: the bargaining strategy of a key potential member of most coalitions delayed the formation of a government until mid-December. Eventually, the leader of New Zealand First, who had bargained with both of the two largest parties (National and Labour), announced on television (without previously informing the other parties' leaders of his decision) that his party would be entering a coalition with National. He was a National party MP until 1994, and a member of the Cabinet, but had campaigned during the general election on the understanding that he would not sustain a further National government. The result of his bargain was to recreate the government which existed before his 'defection', except that he had won promotion for himself to Deputy Prime Minister and Treasurer.[7]

---

[4] He mis-read his notes, according to one commentator (Jackson 1993; Jackson and McRobie 1998)!

[5] See, for example, several of the chapters, including the editor's own, in Holler (1982b).

[6] None of the expansion in parties was the result of by-election victories by previously unrepresented parties.

[7] Labour's unwillingness to appoint him Treasurer over its own candidate was a major reason why its negotiations with the New Zealand First leader (Winston Peters) failed.

3 Qualified Majority Voting in the EU

There was a major conflict within the European Union in 1993 over voting procedures in its Council of Ministers. Although consensus was desirable on most (all?) issues determined there, it was realized that to insist on it would give individual member countries veto rights: there was also substantial unease over employing simple majority voting to determine significant issues. Thus, a compromise had been reached regarding ‘qualified majority voting’ (QMV), whereby a two-thirds majority was needed on certain salient issues: this did not necessarily give one-third of the member countries the ability to block the acceptance of important proposals, however, since the number of votes per country in the Council of Ministers is weighted to reflect population size—though only very roughly so (see Johnston 1994). In the 1980s, as the European Community expanded so the number of votes needed for QMV approval was increased to remain as a constant proportion (0.67) of the total (with the ‘blocking minority’ set at 0.33), but when this was again proposed in 1993 with the likely accession of four more members the British government objected. (In the end only three joined—Austria, Finland and Sweden: the Norwegian electorate rejected their country joining the EU in a referendum.) The size of the ‘blocking minority’ vote (i.e. one-third of the total) should have been increased to 27 following the precedents of previous expansions, but the British government wished to keep it at 23 (0.28 of the total rather than 0.33). In terms of the findings of power index analyses, that opposition seemed counter-productive: the UK government would have been relatively more powerful with a ‘blocking minority’ of 27 rather than 23 (Johnston 1995a, 1995b). The considerable local media coverage of the issue made no reference to that, however; it was assumed that the government was arguing for a voting system that was in the UK’s ‘best interests’.

4 Constitutional Reform in the UK

There is much current discussion of constitutional reform in the United Kingdom, stimulated by pressure groups and the major opposition parties; it includes the possible adoption of an alternative electoral system to fptp (see Bogdanor 1997a). Some steps have been taken along this road: the single transferable vote procedure (STV) is used to elect local councillors, members of the Northern Ireland Assembly, and Members of the European Parliament (MEPs) in Northern Ireland, and also for the election of councillors to Scottish local governments; a non-fptp system (MMP, like that recently-adopted in New Zealand) has been adopted for the Scottish Parliament and the Welsh Assembly established by the newly-elected Labour government after approval was given for both in referendums held in September 1997 (Curtice 1996), as well as for the London Assembly; MEPs from

Great Britain are now elected by a list system of proportional representation[8]; and the Mayor of London and other city mayors are elected by the supplementary vote—a variant of the alternative vote (Hix et al. 2010). Furthermore, the new Labour government elected in 1997 included as one of its manifesto pledges that it would appoint a Commission to study the issue of introducing a more proportional electoral system for the House of Commons, and would follow its report by a referendum at which the electorate could vote for either the status quo (i.e. fptp) or an alternative proposed by the Commission, but the government did not commit itself to recommend a vote for change (see Johnston and Pattie 1997). The report was published in 1998 (Jenkins 1998) but the referendum was never held.

Another issue linked to electoral reform is also widely debated in the UK at present. It is generally referred to as the 'West Lothian question' after the constituency of the MP (Tam Dalyell) who first raised it. If Scotland has its own Parliament, and especially if that has tax-raising as well as spending powers, then should Scottish MPs continue to have votes in the UK House of Commons on issues that relate to areas other than Scotland only? Could this be handled by reducing the number of Scottish MPs in the UK Parliament (where Scotland was already over-represented: see Rossiter et al. 1996)? And how might that be achieved, since the Scots' influence will depend on the distribution of seats among the parties in the House, with all the volatility already described?![9]

5 PR and PP

In all three cases, the many issues that are raised regarding the distribution of power in the relevant bodies have been at best poorly appreciated by the majority of those impacted (or potentially so). Some of those issues related to the details of individual electoral systems (such as those in the German system revealed by Roberts 1996). Others related to wider concerns that are expressed in the simple equation

$$PR \neq PP$$

where *PR* is proportional representation, and *PP* is proportional power.

[8] The MEPs representing Northern Ireland will continue to be elected by stv.

[9] This issue also bedevilled the nineteenth-century British Liberal governments which sought to introduce Home Rule for Ireland within the UK (see Jenkins 1995). After devolution to the Scottish Parliament in 1999 the number of Scottish MPs was reduced from 72 to 59, to achieve near-parity with the electoral quota in England, but this did not eliminate the West Lothian Question; it is still the case that MPs representing Scottish constituencies vote at Westminster on educational legislation relating to England only (and indeed may hold the balance of power in such votes—for example if there is a Labour government with an overall majority in the House of Commons but with fewer MPs elected from English constituencies than its opponents), for example, whereas they cannot influence educational policy in Scotland.

These findings can be summarized (for a more detailed exposition, see Johnston 1998), using parties in a legislative body as the hypothetical example, as:

1. Because a party achieves proportional representation in a legislative body (i.e. its percentage of the seats there is the same as its percentage of the votes cast in the relevant election) this does not mean that it will have the same percentage of the power exercised in the legislature, where power is defined as the number of coalitions which a party can make/break relative to the number in which other parties are involved;
2. The distribution of power is very sensitive to the details of vote disposition across the parties and small changes in that disposition may have substantial effects on the distribution of power (as Holler 1982a, has convincingly demonstrated), perhaps as the result of a by-election or of changes in the party system (fragmentation, as in New Zealand recently, or re-composition as some parties decide to collaborate, if not merge); and
3. It is very difficult to equalize power (i.e. to make each party's percentage of the seats the same as its percentage of the bargaining power). It can be done in some circumstances by determining the size of the majority needed to pass a measure after the composition of the legislature has been determined—but, of course, given the volatility just described, it may be necessary to alter the majority every time the distribution of votes shifts between parties.[10]

These are fundamental findings of much of the work on power indices—and almost irrespective of what index is preferred—but they get little airing during public debates about changing electoral systems.

Several reasons can be suggested for the failure to address these findings in public fora:

1. Those involved in advancing a particular cause, such as the proponents of the STV system in the UK (led by the Electoral Reform Society and the Liberal Democrat party), appear unwilling to address the issues raised here—for them, it seems, *PR* is the universal panacea to all political problems (as illustrated in the discussion after Johnston and Taylor 1985).
2. The implications of many of the findings are either unpalatable to or (in some views at least) perceived as unworkable by those involved in promoting constitutional reform—as with the suggestion of a variable majority figure in a legislature in order to achieve $PR = PP$, or as close to it as possible.[11]

[10] For novel proposals of this method, see Berg and Holler (1985), and Holler (1985). Berg and Holler quote both Buchanan and Tullock (1962) and Curtis (1972) to the effect that a simple majority rule is just one among a wide range of possibilities—and it can maximize the expected number of disappointed voters.

[11] Berg and Holler (1985, p. 428) accept, for example, that a randomized decision procedure can compromise the stability and continuity of a decision-making process.

3. The research findings have not been brought to the attention of those involved in the public arena, in a form which they can appreciate and transmit[12]; and
4. The academics who conduct the research are not particularly interested in making their findings ‘applicable’, either in the technocratic sense of giving the designers of electoral systems tools that they can use or in the emancipatory sense of educating citizens regarding the full nature of the systems they are called upon to operate and perhaps replace by others.[13]

Whatever the reason (or combination of reasons) in any particular situation, the result is invariably a poorly-informed population, and almost certainly a partially-informed political elite too. Thus, many discussions of electoral reform in the UK are characterized by a debate between, on the one hand, advocates of *PR* who present what they identify as a moral case without any consideration of the consequences of a Parliament elected by a *PR* procedure and, on the other hand, those who say that *PR* stimulates a need for coalition government, which is undemocratically negotiated in secret by the parties: the switch of the German FDP from sustaining an SPD government to a CDU-CSU one mid-term is frequently quoted as the quintessential exemplar of this, along with the—falsely claimed (see Donovan 1996)—instability of coalitions in Italy. In discussing constitutional reform, for example, Tony Blair (leader of the British Labour party and Prime Minister of the UK 1997–2007) claimed in 1996 that:

> I personally remain unpersuaded that proportional representation would be beneficial for the Commons. It is not, as some claim, a simple question of moving from an “unfair” to a “fair” voting system. A electoral system must meet two democratic tests: it needs to reflect opinion, but it must also aggregate opinion without giving disproportionate influence to splinter groups. Aggregation is particularly important for a parliament whose job is to create and sustain a single, mainstream government (Blair 1996, p. 35).[14]

This continues the frequently-heard case in the UK for strong majority government, which leads to a defence of an electoral system in which no government elected since the Second World War has obtained the support of a majority of the electorate and most have been sustained by a ‘manufactured majority’—a majority

---

[12] Two major exceptions to this were the Commission established by the Labour party in the UK to consider the issues of electoral reform, which brought together a great deal of research into all aspects of the subject (Plant 1991, 1993), and the Royal Commission established by the New Zealand government in the mid-1980s, on whose detailed research the eventual adoption of electoral reform there was based (Royal Commission 1986): in neither case were the issues raised here extensively debated, however, and they certainly were not part of the wider debate that followed the Commissions’ reports.

[13] For an exception to this in the UK, see Dunleavy et al. (1998).

[14] Bogdanor (1997b, p. 80) also set two tests for an electoral system: ‘to ensure that the majority rules’ and ‘to ensure that all significant minorities are represented’. He concluded that electoral reform—specifically stv—is needed to meet these two, whereas Tony Blair concluded that fptp best meets his two criteria. For a response to Bogdanor, see Norton (1997).

of the seats based on a minority of the votes cast.[15] The argument could have been made in the context of the three main findings of power index studies listed above, but it was not. If it had been, then the discussion of the way forward might have involved advocates of *PR* accepting Blair's point regarding splinter groups, and promoting discussion of alternative constitutional changes, such as:

- The variable majority solution identified earlier;
- Separation of the legislature from the executive (as in France and the United States), allowing the electorate to make clear and distinct separate statements regarding who they wanted to govern them and how they wanted that government to be checked in the legislature (as in the USA in 1996, which elected a Democratic President but returned Republican majorities in both Houses of Congress and with the recent periods of co-habitation between a President and a Prime Minister of different ideological persuasions in France); or
- A federal system which has a 'people's house' elected on PR and a 'state's house' elected to give either equal or equitable representation to the country's major subdivisions.

Discussion of these is largely notable by its absence.

6 Conclusions

In sum, academics have done a great deal of theoretical and quantitative empirical research exploring the nature of power and illuminating, for themselves at least, its complexities and inherent difficulties. Many of their findings raise difficulties that those who operate political and electoral systems find extremely troublesome, or would do if they were aware of them. Unfortunately, the academics involved have not disseminated their findings widely among such people, let alone among the general population.[16] As a result, debates about the construction of electoral and political systems in some countries (notably eastern Europe in recent years: see Benoit 1996, on Hungary) and about constitutional change in others have been

[15] The government formed after the 2010 general election when no party gained a majority of seats is the country's first 'peacetime' coalition since 1900 in which the government has a majority of both the votes and the seats in the House of Commons (the Conservatives gained 36.1 % of the votes and their coalition partners the Liberal Democrats 23.0 %; they obtained 47.0 and 8.8 % of the seats respectively). One issue that has since arisen is relative power within the government: some Conservative MPs claim that with 5 of the 21 Cabinet seats the Liberal Democrats are over-represented and -powerful there; the Liberal Democrats also occupy 16 sub-Cabinet ministerial posts and in total hold some 16 % of all of the official government posts in both the House of Commons and the House of Lords. Their power appears to Conservative critics of the coalition to be incommensurate with their share of the seats in the House of Commons (8.7 % of the total but 15.7 % of all those occupied by one of the coalition parties!).

[16] There are exceptions, including several in the Eastern European countries which recently experienced a transition back to democracy.

conducted in almost total ignorance of the research findings, usually employing gross political slogans that have little basis in 'fact'. Thus, those who eventually determine the merits of the cases presented, the electorate, are denied the sort of sophisticated information which, properly expressed, could inform their decision-making. This situation presents a substantial challenge to us as researchers[17]: whether we see our roles as technocrats or as emancipators we should ensure that our research informs debates about the structures within which political life is played-out.

Acknowledgments I am grateful to Charles Pattie for insightful comments on a draft of the first published version of this polemic.

Postscript 2011

The 2010 general election in the United Kingdom resulted in no party having a majority in the House of Commons, so for the first time (other than during the two twentieth century world wars) the country had a coalition government. During the campaign preceding that election, the opinion polls suggested that no party would win an overall majority and commentators argued that, because of the biases currently inherent to the UK electoral system (Johnston et al. 2001), the allocation of seats across the three main parties could vary considerably with small differences in their share of the votes (see also the simulations in Rallings and Thrasher 2007). The issue of which parties might be able to form a viable coalition was much discussed but there was, however, no formal modeling of each party's relative power in that evolving situation.

One of the coalition government's first major pieces of legislation proposed that the voting system for the House of Commons be changed from first-past-the-post to the alternative vote, with the decision to be made by the electorate in a binding referendum (Johnston and Pattie 2011). There was much discussion that this would almost certainly mean the end of one-party government and its replacement by coalitions and a great deal of the debate regarding this focused on the power it would give to relatively small parties, probably of the centre (as with the Free Democrats for long periods in post-1945 Germany)—but there was no formal modeling. Furthermore, there was little recognition that the House of Commons now includes MPs from eight parties with more than one member (apart from the Conservatives, Labour and Liberal Democrats these are: Plaid Cymru—the Party of Wales; the Scottish National Party; and, from Northern Ireland, the Democratic Unionist Party, the Social and Democratic Labour Party, and Sinn Féin) plus three

[17] Meeting that challenge is not easy: several attempts that I have made (Johnston 1982, 1995a; Johnston and Taylor 1985) to stimulate interest among (a) those committed to electoral reform in the UK and (b) students of politics have almost entirely failed. I had greater hopes for Johnston and Pattie (1997), but nothing changed!

other individuals (representatives of the Green Party, the Alliance Party of Northern Ireland, and an independent—also from Northern Ireland). Those smaller parties might exercise considerable power, both individually and severally (although the five Sinn Féin MPs do not take their seats and vote), and modeling could have illustrated a range of such scenarios—including those in which parties such as Plaid Cymru, the Scottish National Party and the Green Party might have increased their representation in a Parliament elected by AV (though see Sanders et al. 2011). The potential for modeling that could illuminate the likely situations in what was an ill-informed and bad-tempered referendum campaign in 2011 was not realised, however. The referendum failed, with the switch to AV being rejected by a ratio of 68:32.

The UK Parliament is now discussing whether to replace the appointed House of Lords by an elected second chamber (with probably 20 % of the members still appointed, by a non-partisan commission). The current House of Lords comprises four main groupings—those who take the three main party whips (Conservative, Labour, Liberal Democrat) plus the cross-benchers (who have no party affiliations); there is a small number of 'others', including 26 bishops of the Church of England. None of the parties has a majority although the current coalition of the Conservative and Liberal Democrat parties has a majority among those who take a party whip; whether it can get its business through the House thus depends on the votes of the cross-benchers (many of whom attend the House only infrequently: Johnston and Pattie 2011). The current proposal is for the elected members of a revised House to be chosen by the quasi-proportional STV method from multi-member constituencies; the equivalent to the non-partisan cross-benchers would be appointed but, unlike their predecessors, expected to be full-time members. It is thus very unlikely that any one party—or even a coalition of two—would have a majority in such a House, and the cross-benchers (and/or other small groups outwith the main parties) could be crucial to the fate of any legislation,[18] depending on how they split and assuming that members of each party group all vote the same way. This offers a major opportunity for modelling the power not only of the party groups but also, more interestingly, of the non-party cross-benchers, individually and in combinations.

The New Zealand case is a further example of the absence of any formal modelling of the allocation of power in legislatures where no party has a majority of the seats. Following the decision to change the electoral system to MMP in 1993 the country has experienced a range of single-party and multi-party governments, and it is that experience—without any detailed analysis using power indices—that informed the electorate when it was asked in a 2011 referendum to decide whether

[18] In a quasi-PR system parties that fail to win seats at general elections—notably the United Kingdom Independence Party (UKIP) which won 16.5 % of the votes at the 2009 European Parliament elections (which used a closed list PR electoral system) and 13 of the 72 seats (McLean and Johnston 2010)—may gain representation in an elected House of Lords. To try and prevent this, the proposal is for election to be held for one-third of the House every five years in 6–7 member constituencies, on the same day as elections to the House of Commons.

to retain MMP or switch to another system. The referendum result was not binding on the government. If the decision was for change, then Parliament would decide whether to hold a further referendum in 2014 when the electorate would choose between MMP and the most popular of the four other systems voted on in the second part of the 2011 referendum. (Even if the electors opted to retain MMP in 2011 the referendum ballot included a second question asking which of four other systems they would prefer if MMP was to be replaced.) If the electorate voted to keep MMP then there would be an independent review of that system in 2012 to determine if any elements of it should be changed; in the event, the country voted by 58:42 to retain MMP, with fptp the most popular among the four options offered in the second part of the referendum.[19] Much was done to educate the electorate regarding the various electoral systems on the referendum menu, but little on the possible consequences in terms of the allocation of power.[20]

Finally, one example where academics have been involved as technocrats, using their research expertise in the measurement of power, concerns the allocation of seats in the European Parliament. A group of mathematicians recommended that each member state be allocated a baseline of five seats and that the remaining seats be allocated proportional to the states' populations so that: (1) no state has more seats than a larger state; and (2) the ratio of population to seats increases with state population—a principle known as 'degressive proportionality' (Grimmett 2011). This is a rare example of theoretical knowledge in this field being put to practical effect (see also Rose and Bernhagen 2010; Rose et al. 2012). Some academics have designed electoral systems that they consider superior to those either currently being deployed or considered (e.g. Brams 2008; Balinski and Laraki 2011—see also Szpiro 2010); they have had some success, notably in the design of systems for local governments (e.g. Balinski and Ramirez Gonzales 1999; Schuster et al. 2003; Pukelsheim 2006, 2009) but as yet not for elections to national legislatures.

References

Bechofer, F. (1996). Quantitative research in British sociology: has it changed since 1981? *Sociology, 30*, 583–591.

Benoit, K. (1996). Hungary's 'two-vote' electoral system. *Representation, 33*, 162–170.

[19] It got 31 % of the votes with 8 % for preferential voting (i.e. the alternative vote), 11 % for STV, and 16 % for the supplementary vote system (mixed member majoritarian—like MMP but the list vote is not used to compensate for disproportionality in the results of the constituency contests).

[20] In the main leaflet sent to all voters, the sections on fptp and preferential voting included a final sentence that 'A government can usually be formed without the need for coalitions or agreements between parties' whereas those for MMP, stv and the supplementary member included 'Coalitions or agreements between political parties are usually needed before governments can be formed'.

Berg, S., & Holler, M. J. (1985). Randomized decision rules in voting games: a model for strict proportional power. *Quality and Quantity, 20*, 419–429.

Blair, T. (1996). Blair on the constitution: Democracy's second age. *The Economist*, (7983), 33–35.

Bogdanor, V. (1997a). *Power to the people: a guide to constitutional reform*. London: Victor Gollancz.

Bogdanor, V. (1997b). First-past-the post: an electoral system which is difficult to defend. *Representation, 34*, 80–83.

Boston, J., Levine, S., McLeay, E., & Roberts, N. S. (1996). Why did New Zealand adopt German-style proportional representation? *Representation, 33*, 134–140.

Brams, S. J., & Fishburn, P. C. (1983). *Approval voting*. Boston: Birkhauser.

Buchanan, J. M., & Tullock, G. (1962). *The calculus of consent: The logical foundations of constitutional democracy*. Ann Arbor: The University of Michigan Press.

Catt, H., Harriss, P., & Roberts, N. S. (1992). *Voter's Choice: Electoral change in New Zealand*. Palmerston North: Dunmore Press.

Curtice, J. (1996). Why the additional member system has won out in Scotland. *Representation, 33*, 119–124.

Curtis, R. B. (1972). Decision-rules and collective values in constitutional choice. In R. G. Niemi & H. F. Weisberg (Eds.), *Probability models of collective decision-making*. Columbus: Charles Merrill.

Donovan, M. (1996). Electoral reform in Italy. *Representation, 33*, 141–147.

Dunleavy, P., Margetts, H., O'Duffy, B. & Weir, S. (1998). *Making Votes Count*. Colchester: Democratic Audit.

Hix, S., Johnston, R. J., & McLean, I. (2010). *Choosing an electoral system*. London: The British Academy.

Holler, M. J. (1982a). Note on a paradox. *Jahrbuch für Nationalökonomie und Statistik, 197*, 251–257.

Holler, M. J. (Ed.). (1982b). *Power, voting and voting power*. Würzburg-Wien: Physica-Verlag.

Holler, M. J. (1985). Strict proportional power in voting bodies. *Theory and Decision, 23*, 249–258.

Jackson, K. (1993). The origins of the electoral referendums. In A. D. McRobie (Ed.), *Taking it to the people? The New Zealand Electoral Reform Debate* (pp. 12–23). Christchurch: Hazard Books.

Jackson, K., & McRobie, A. D. (1998). *New Zealand adopts proportional representation: Accident? Design? Evolution?*. Aldershot: Ashgate.

Jenkins, R. (1995). *Gladstone*. London: Macmillan.

Jenkins, R. (1998). *The report of the independent commission on the voting system*. London: Stationery Office.

Johnston, R. J. (1982). An unresolved issue for electoral reformers. *Representation, 23*(90), 6–10.

Johnston, R. J. (1994). The conflict over voting power in the European Union Council of Ministers. *Environment and Planning A, 26*, 1001–1006.

Johnston, R. J. (1995a). Proportional representation and proportional power. *Politics Review, 4*(4), 28–33.

Johnston, R. J. (1995b). The conflict over qualified majority voting in the European Union Council of Ministers: An analysis of the UK negotiating stance using power indices. *British Journal of Political Science, 25*, 245–253.

Johnston, R. J. (1998). Proportional representation and a 'fair electoral system' for the United Kingdom. *Journal of Legislative Studies, 4*, 128–148.

Johnston, R. J., & Pattie, C. J. (1997). Electoral reform without constitutional reform. *The Political Quarterly, 68*, 379–387.

Johnston, R. J., & Taylor, P. J. (1985). People, places and parliaments: a geographical perspective on electoral reform in Great Britain. *The Geographical Journal, 151*, 327–338.

Morris, P. (1996). Qualified majority voting and power indices: a further response to Johnston. *British Journal of political Science, 26*, 595–598.

Norton, P. (1997). The case for first-past-the-post. *Representation, 34*, 84–88.

Plant, L. (1991). *"The plant report: A working party on electoral reform", Guardian Studies, 3*. London: The Guardian newspaper.

Plant, L. (1993). *Report of the working party on electoral systems*. London: The Labour Party.

Roberts, G. K. (1996). Neglected aspects of the German electoral system. *Representation, 33*, 125–133.

Riba, C. (1996). The use of mathematics in political science: a survey of European and American journals. *European Journal of Political Research, 29*, 477–508.

Rossiter, D. J., Pattie, C. J., & Johnston, R. J. (1996). New boundaries, old inequalities: The evolution and partisan impact of the Celtic preference in British redistriction. In I. Hampsher-Monk & J. Stanyer (Eds.), *Contemporary political studies* (pp. 1469–1482). Belfast: Political Studies Association.

Royal Commission on the Electoral System. (1986). *Towards a better democracy*. Wellington: Government Printer.

References in Postscript

Balinski, M., & Laraki, R. (2011). *Majority judgment: Measuring, Ranking and Electing*. Cambridge: The MIT Press.

Balinski, M., & Ramirez Gonzales, V. (1999). Parametric methods of apportionment, rounding and production. *Mathematical Social Sciences, 37*, 107–122.

Brams, S. J. (2008). *Mathematics and democracy: Designing better voting and fair-division procedures*. Princeton: Princeton University Press.

Grimmett, G. et al., (2011). *The Allocation between the EU member states of seats in the European Parliament*, Brussels, The European Parliament, Directorate-General for Internal Policies, PE 432.760.

Johnston, R. J., & Pattie, C. J. (2011). Parties and cross-benchers voting in the post-2010 house of lords: the example of the parliamentary voting system and Constituencies bill. *British Politics, 6*, 430–452.

Johnston, R. J., Pattie, C. J., Dorling, D., & Rossiter, D. J. (2001). *From votes to seats: The operation of the UK electoral system since 1945*. Manchester: Manchester University Press.

McLean, I., & Johnston, R. J. (2010). When is proportional not proportional? Great Britain's 2009 elections to the European Parliament. *Representation, 45*, 349–355.

Pukelsheim, F. (2006). Current issues of apportionment methods. Available online at http://www.math.uni-augsburg.de/stochastik/pukelsheim/2006c.pdf.

Pukelsheim, F. (2009). Zurich's new apportionment. *German Research, 31*, 10–12.

Rallings, C., & Thrasher, M. (2007). *Media guide to the new Parliamentary Constituencies*. Plymouth: Local Government Chronicle Elections Centre.

Rose, R., & Bernhagen, P. (2010). Inequality in the representation of citizens in the European Parliament. Abderdeen: University of Aberdeen, *Studies in Public Policy* 472.

Rose, R., Bernhagen, P. & Broz, G. (2012) Evaluating competing criteria for allocating Parliamentary seats. *Journal of Mathematical Social Sciences*, *63*, 85–89.

Sanders, D., Clarke, H. D., Stewart, M. C., & Whiteley, P. (2011). Simulating the effects of the alternative vote in the 2010 general election. *Parliamentary Affairs, 64*, 5–23.

Schuster, K., Pukelsheim, F., Drton, M., & Draper, N. R. (2003). Seat biases of apportionment for proportional representation. *Electoral Studies, 22*, 651–676.

Szpiro, G. (2010). *Numbers rule: The vexing mathematics of democracy from Plato to the present*. Princeton NJ: Princeton University Press.

Fair Voting Rules in Committees

František Turnovec

1 Introduction

Let us consider a committee with n members. Each member has some voting weight (number of votes, shares etc.,) and a voting rule is defined by a minimal number of weights required for passing a proposal. Given a voting rule, voting weights provide committee members with voting power. Voting power means an ability to influence the outcome of voting. Voting power indices are used to quantify the voting power.

The concept of fairness is being discussed related to the distribution of voting power among different actors of voting. This problem was clearly formulated by Nurmi (1982, p. 204): "If one aims at designing collective decision-making bodies which are democratic in the sense of reflecting the popular support in terms of the voting power, we need indices of the latter which enable us to calculate for any given distribution of support and for any decision rule the distribution of seats that is 'just'. Alternatively, we may want to design decision rules that—given the distribution of seats and support—lead to a distribution of voting power which is identical with the distribution of support".

Voting power is not directly observable; voting weights are used as a proxy. Therefore, fairness is usually defined in terms of voting weights (e.g., voting weights are proportional to the results of an election). Assuming that a principle of fair distribution of voting weights is selected, we are addressing the question of how to achieve equality of voting power (at least approximately) to voting weights. The concepts of strict proportional power and the randomized decision

Earlier version of this chapter has been published in *Homo Oeconomicus* 27(4); see Turnovec (2011).

F. Turnovec (✉)
Institute of Economic Studies, Charles University in Prague, Opletalova 26, 110 00 Prague 1, Czech Republic
e-mail: turnovec@fsv.cuni.cz

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*, DOI: 10.1007/978-3-642-35929-3_19, © Springer-Verlag Berlin Heidelberg 2013

rule introduced by Holler (1982a, 1985, 1987), of optimal quota of Słomczyński and Życzkowski (2007), and of intervals of stable power (Turnovec 2008b) are used to find, given voting weights, a voting rule minimizing the distance between actors' voting weights and their voting power.

Concept of fairness is frequently associated with so-called square root rule, attributed to British statistician Lionel Penrose (1946). The square root rule is closely related to indirect voting power measured by the Penrose-Banzhaf power index.[1] Different aspects of the square root rule have been analysed in Felsenthal and Machover (1998, 2004), Laruelle and Widgrén (1998), Baldwin and Widgrén (2004), Turnovec (2009). The square root rule of "fairness" in the EU Council of Ministers voting was discussed and evaluated in Felsenthal and Machover (2007), Słomczyński and Życzkowski (2006, 2007), Hosli (2008), Leech and Aziz (2008), Turnovec (2008a) and others. Nurmi (1997a) used this rule to evaluate the representation of voters' groups in the European Parliament.

Section 2 introduces basic definitions and shortly resumes the applied power indices methodology. Section 3 introduces the concept of quota intervals of stable power and optimal quota. Section 4 applies the concept of optimal quota (fair voting rule) on the Lower House of the Czech Parliament. While the framework of the analysis of fairness is usually restricted to the Penrose-Banzhaf concept of power, we are treating it in a more general setting and our results are relevant for any power index based on pivots or swings and for any concept of fairness.

2 Committees and Voting Power

A simple weighted committee is a pair [N, **w**], where N will be a finite set of n committee members $i = 1, 2, \ldots, n$, and $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ will be a nonnegative vector of committee members' voting weights (e.g., votes or shares). By 2^N we denote the power set of N (set of all subsets of N). By a voting coalition we mean an element $S \in 2^N$, i.e., a subset of committee members voting either YES or NO. $w(S) = \sum_{i \in S} w_i$ denotes the voting weight of coalition S.

[1] The square root rule is based on the following propositions: Let us assume n units with population $p_1, p_2, \ldots, p_n$, and the system of representation by a super-unit committee with voting weights $w_1, w_2, \ldots, w_n$. It can be rigorously proved that for sufficiently large p_i the absolute Penrose-Banzhaf power of individual citizen of unit i in unit's referendum is proportional to the square root of p_i. If the relative Penrose-Banzhaf voting power of unit i representation is proportional to its voting weight, then indirect voting power of each individual citizen of unit i is proportional to the product of voting weight w_i and square root of population p_i. Based on the conjecture (not rigorously proved) that for n large enough the relative voting power is proportional to the voting weights, the square root rule concludes that the voting weights of the units' representations in the super-unit committee, proportional to square roots of units' population, lead to the same indirect voting power of each citizen independently of the unit she is affiliated with.

The voting rule is defined by quota q satisfying $0 < q \leq w(N)$, where q represents the minimal total weight necessary to approve the proposal. Triple [N, q, **w**] we call a *simple quota weighted committee*. The voting coalition S in committee [N, q, **w**] is called a winning one if $w(S) \geq q$ and a losing one in the opposite case. The winning voting coalition S is called critical if there exists at least one member $k \in S$ such that $w(S \backslash k) < q$ (we say that k is critical in S). The winning voting coalition S is called minimal if any of its members is critical in S.

A priori voting power analysis seeks an answer to the following question: Given a simple quota weighted committee $[N, q, \mathbf{w}]$, what is an influence of its members over the outcome of voting? The absolute voting power of a member i is defined as a probability $\Pi_i[N, q, \mathbf{w}]$that i will be decisive in the sense that such a situation appears in which she would be able to decide the outcome of voting by her vote (Nurmi 1997b and Turnovec 1997). The corresponding relative voting power is defined as

$$\pi_i[N, q, w] = \frac{\Pi_i[N, q, w]}{\sum\limits_{k \in N} \Pi_k[N, q, w]}$$

Three basic concepts of decisiveness are used: swing position, pivotal position and membership in a minimal winning coalition (MWC position). The s*wing position* is an ability of an individual voter to change the outcome of voting by a unilateral switch from YES to NO. (If member j is critical with respect to a coalition S, we say that he has a swing in S.) The *pivotal position is* such a position of an individual voter in a permutation of voters expressing a ranking of attitudes of members to the voted issue (from the most preferable to the least preferable) and the corresponding order of forming of the winning coalition, in which her vote YES means a YES outcome of voting and her vote NO means a NO outcome of voting. (We say that j is pivotal in the permutation considered.) The MWC position is an ability of an individual voter to contribute to a minimal winning coalition (membership in the minimal winning coalition).

Let us denote by W(N, q, **w**) the set of all winning coalitions and by W_i(N, q, **w**) the set of all winning coalitions with i, C(N, q, **w**) as the set of all critical winning coalitions, and by C_i(N, q, **w**) the set of all critical winning coalitions i has the swing in, by P(N, q, **w**) the set of all permutations of N and P_i(N, q, **w**), the set of all permutations i is pivotal in, M(N, q, **w**) the set of all minimal winning coalitions, and M_i(N, q, w) the set of all minimal winning coalitions with i. By card(S) we denote the cardinality of S; of course, card($\emptyset$) = 0.

Assuming many voting acts and all coalitions equally likely, it makes sense to evaluate the a priori voting power of each member of the committee by the probability to have a swing, measured by the absolute Penrose-Banzhaf (PB) power index (Penrose 1946; Banzhaf 1965):

$$\Pi_i^{PB}(N, q, w) = \frac{card(C_i)}{2^{n-1}}$$

Here card(C_i) is the number of all winning coalitions the member i has the swing in and 2^{n-1} is the number of all possible coalitions with i as a member. To compare the relative power of different committee members, the relative form of the PB power index is used:

$$\pi_i^{PB}(N,q,w) = \frac{card(C_i)}{\sum_{k \in N} card(C_k)}$$

While the absolute PB is based on a well-established probability model (see e.g., Owen 1972), its normalization (relative PB index) destroys this probabilistic interpretation, the relative PB index simply answers the question of what is voter i's share in all possible swings.

Assuming many voting acts and all possible preference orderings equally likely, it makes sense to evaluate an a priori voting power of each committee member by the probability of being in pivotal situation, measured by the Shaply-Shubik (SS) power index (Shapley and Shubik 1954):

$$\Pi_i^{SS}(N,q,w) = \frac{card(P_i)}{n!}$$

Here card(P_i) is the number of all permutations in which the committee member i is pivotal, and n! is the number of all possible permutations of committee members). Since $\sum_{i \in N} card(P_i) = n!$ it holds that

$$\pi_i^{SS}(N,q,\mathbf{w}) = \frac{card(P_i)}{\sum_{k \in N} card(P_k)} = \frac{card(P_i)}{n!}$$

i.e., the absolute and relative form of the SS-power index is the same.[2]

Assuming many voting acts and all possible coalitions equally likely, it makes sense to evaluate the voting power of each committee member by the probability of membership in a minimal winning coalition, measured by the absolute Holler-Packel (HP) power index

[2] Supporters of the Penrose-Banzhaf power concept sometimes reject the Shapley-Shubik index as a measure of voting power. Their objections to the Shapley-Shubik power concept are based on the classification of power measures on so-called I-power (voter's potential influence over the outcome of voting) and P power (expected relative share in a fixed prize available to the winning group of committee members, based on cooperative game theory) introduced by Felsenthal, Machover and Zwicker (1998). The Shapley-Shubik power index was declared to represent P-power and as such is unusable for measuring influence in voting. We tried to show in Turnovec (2007) and Turnovec, Mercik, Mazurkiewicz (2008) that objections against the Shapley-Shubik power index, based on its interpretation as a P-power concept, are not sufficiently justified. Both Shapley-Shubik and Penrose-Banzhaf measure could be successfully derived as cooperative game values, and at the same time both of them can be interpreted as probabilities of being in some decisive position (pivot, swing) without using cooperative game theory at all.

$$\Pi_i^{HP}(N, q, \mathbf{w}) = \frac{card(M_i)}{2^n}$$

Here card(M_i) is the number of all minimal winning coalitions with i, and 2^n is the number of all possible coalitions).[3] Originally the HP index was defined and is usually being presented in its relative form (Holler 1982b; Holler and Packel 1983), i.e.,

$$\pi_i^{HP}(N, q, \mathbf{w}) = \frac{card(M_i)}{\sum_{k \in N} card(M_k)}$$

The above definition of the absolute HP index allows a clear probabilistic interpretation. Multiplying and dividing it by the card (M), we obtain

$$\Pi_i^{HP}(N, q, \mathbf{w}) = \frac{card(M_i)}{card(M)} \frac{card(M)}{2^n}$$

In this breakdown the first term gives the probability of being a member of a minimal winning coalition, provided the MWC is formed, and the second term the probability of forming a minimal winning coalition assuming that all voting coalitions are equally likely. The relative HP index has the same problem with a probabilistic interpretation as the relative PB index.[4]

In the literature there are still two other prominent concepts of power indices: the Johnston (J) power index based on swings, and the Deagan-Packel (DP) power index, based on membership in minimal winning coalitions. The Johnston power index (Johnston 1978) measures the power of a member of a committee as a normalized weighted average of the number of her swings, using as weights the reciprocals of the total number of swings in each critical winning coalition. (The swing members of the same winning coalition have the same power, which is equal to 1/[# of swing members]). The Deegan-Packel power index (Deegan and Packel 1978) measures the power of a member of a committee as a normalized weighted average of the number of minimal critical winning coalitions he is a member of, using as weights the reciprocals of the number of players in a minimal winning coalition.

It is difficult to provide some intuitively acceptable probabilistic interpretation for relative Johnston index and Deegan-Peckel index. They provide a normative scheme of the division of rents in the committee rather than a measure of an a priori power. (In the sense of Felsenthal and Machover (1998) classification they can be considered as measures of P power).

---

3 The definition of an absolute HP power index is provided by the author (a similar definition of absolute PB power can be found in Brueckner (2001), the only difference is that we relate the number of MWC positions of member i to the total number of coalitions, not to the number of coalitions of which i is a member).

4 For a discussion about the possible probabilistic interpretation of the relative PB and HP, see Widgrén (2001).

It can be easily seen that for any $\alpha > 0$ and any power index based on swings, pivots or MWC positions it holds that $\Pi_i[N, \alpha q, \alpha \mathbf{w}] = \Pi_i[N, q, \mathbf{w}]$. Therefore, without the loss of generality, we shall assume throughout the text that $\sum_{i \in N} w_i = 1$ and $0 < q \leq 1$, using only relative weights and relative quotas in the analysis.

3 Quota Interval of Stable Power, Fairness and Optimal Quota

Let us formally define a few concepts we shall use later in this chapter.

Definition 1 A simple weighted committee [N, **w**] has a property of *strict proportional power* with respect to a power index $\boldsymbol{\pi}$, if there exists a voting rule q* such that $\boldsymbol{\pi}[N, q*, \mathbf{w}] = \mathbf{w}$, i.e., the relative voting power of committee members is equal to their relative voting weights.

In general, there is no reason to expect that such a voting rule exists. However, the concepts of randomized voting rule and strict proportional expected power were introduced by Holler (1982a, 1985), and studied by Berg and Holler (1986).

Definition 2 Let [N, **w**] be a simple weighted committee, $\mathbf{q} = (q_1, q_2, \ldots, q_m)$ be a vector of different quotas, $\boldsymbol{\pi}^k$ be a relative power index for quota q_k, and $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_m)$ be a probability distribution over elements of **q**. The *randomized voting rule* (**q**, $\boldsymbol{\lambda}$) selects within different voting acts by random mechanism quotas from **q** by the probability distribution $\boldsymbol{\lambda}$. Then [N, **w**] has a property of *strict proportional expected power* with respect to a relative power index $\boldsymbol{\pi}$, if there exists a randomized voting rule (**q***, $\boldsymbol{\lambda}$*) such that the vector of the mathematical expectations of power is equal to the vector of voting weights:

$$\boldsymbol{\pi}(N, (\mathbf{q}, \boldsymbol{\lambda}), \mathbf{w}) = \sum_{k=1}^{m} \lambda_k \boldsymbol{\pi}^k(N, q_k, \mathbf{w}) = \mathbf{w}$$

Definition 3 Let $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ be a fair distribution of voting weights (with whatever principle is used to justify it) in a simple weighted committee [N, **w**], $\boldsymbol{\pi}$ is a relative power index, ($\boldsymbol{\pi}$[N, q, **w**] is a vector valued function of q), and d is a distance function, then the voting rule q_1 is said to be *at least as fair* as voting rule q_2 with respect to the selected $\boldsymbol{\pi}$ if $d(\mathbf{w}, \boldsymbol{\pi}(N, q_1, \mathbf{w})) \leq d(\mathbf{w}, \boldsymbol{\pi}(N, q_2, \mathbf{w}))$.

Intuitively, given **w**, the voting rule q_1 is preferred to voting rule q_2 if q_1 generates a distribution of power closer to the distribution of weights than q_2.

Definition 4 The voting rule q* that minimizes a distance d between $\boldsymbol{\pi}$[N, q, **w**] and **w** is called an optimal voting rule (*optimal quota).*with respect to the selected power index π.

Let [N, q, **w**] be a simple weighted quota committee and C_{is} be the set of critical winning coalitions of the size s in which i has a swing, then

$$card(P_i) = \sum_{s \in N} card(C_{is})(s-1)!(n-s)!$$

is the number of permutations with the pivotal position of i in [N, q, **w**]. The number of pivotal positions corresponds to the number and structure of swings. If in two different committees sets of swing coalitions are identical, then the sets of pivotal positions are also the same.

Proposition 1 *Let* $[N, q_1, \mathbf{w}]$ *and* $[N, q_2, \mathbf{w}]$, $q_1 \neq q_2$, *be two simple quota-weighted committees such that* $W[N, q_1, \mathbf{w}] = W[N, q_2, \mathbf{w}]$, *then*

$$C_i(N, q_1, \mathbf{w}) = C_i(N, q_2, \mathbf{w})$$
$$P_i(N, q_1, \mathbf{w}) = P_i(N, q_2, \mathbf{w})$$

and

$$M_i(N, q_1, \mathbf{w}) = M_i(N, q_2, \mathbf{w})$$

for all $i \in N$.

From Proposition 1 it follows that in two different committees with the same set of members, the same weights and the same sets of winning coalitions, the PB-power indices, SS-power indices and HP-power indices are the same in both committees, independently of quotas. Moreover, since the Johnston index is based on the concept of swing and the Deegan-Packel power index is based on membership in minimal winning coalitions, the two indices give the same.

Proposition 2 *Let* [N, q, **w**] *be a simple quota weighted committee with a quota q,*

$$\mu^+(q) = \min_{S \in W[N,q,w]} (w(S) - q)$$

and

$$\mu^-(q) = \min_{S \in 2^N \setminus W(N,q,w)} (q - w(S))$$

Then for any particular quota q we have $W[N, q, \mathbf{w}] = W[N, \gamma, \mathbf{w}]$ *for all* $\gamma \in (q-\mu^-(q), q + \mu^+(q)]$.

Proof

(a) Let $S \in W[N, q, \mathbf{w}]$, then from the definition of $\mu^+(q)$

$$w(S) - q \geq \mu^+(q) \geq 0 \Rightarrow w(S) - q - \mu^+(q) \geq 0 \Rightarrow S \in W(N, q + \mu^+, w),$$

hence S is winning for quota $q + \mu^+(q)$. If S is winning for $q + \mu^+(q)$, then it is winning for any quota $\gamma \leq q + \mu^+(q)$.

(b) Let $S \in 2^N \setminus W[N, q, \mathbf{w}]$, then from the definition of $\mu^-(q)$

$$q - w(S) \geq \mu^-(q) \geq 0 \Rightarrow q - \mu^-(q) - w(s) \geq 0 \Rightarrow S \in 2^N \setminus W(N, q - \mu^-(q), w),$$

hence S is losing for quota $q - \mu^-(q)$. If S is losing for $q - \mu^-(q)$, then it is losing for any quota $\gamma \geq q - \mu^-(q)$.

From (a) and (b) it follows that for any $\gamma \in (q - \mu^-(q), (q - \mu^+(q)]$

$$S \in W(N, q, \mathbf{w}) \Rightarrow S \in W(N, \gamma, \mathbf{w})$$
$$S \in \{2^N \backslash W(N, \gamma, \mathbf{w})\} \Rightarrow S \in \{2^N \backslash W(N, q, \mathbf{w})\}$$

which implies that $W(N, q, \mathbf{w}) = W(N, \gamma, \mathbf{w})$. □

From Propositions 1 and 2 it follows that swing, pivot and MWC-based power indices are the same for all quotas $\gamma \in (q - \mu^-(q), q + \mu^+(q)]$. Therefore, the interval of quotas $(q - \mu^-(q), q + \mu^+(q)]$ we call an *interval of stable power* for quota q. Quota $\gamma^* \in (q - \mu^-(q), q + \mu^+(q)]$ is called the marginal quota for q if $\mu^+(\gamma^*) = 0$.

Now we define a partition of the power set 2^N into equal weight classes $\Omega_0, \Omega_1, \ldots, \Omega_r$ (such that the weight of different coalitions from the same class is the same and the weights of different coalitions from different classes are different). For the completeness set $w(\emptyset) = 0$. Consider the weight-increasing ordering of equal weight classes $\Omega^{(0)}, \Omega(1), \ldots, \Omega^{(r)}$ such that for any $t < k$ and $S \in \Omega(t)$, $R \in \Omega(k)$ it holds that $w(S) < w(R)$. Denote $q_t = w(S)$ for any $S \in \Omega^{(t)}$, $t = 1, 2, \ldots, r$.

Proposition 3 *Let $\Omega^{(0)}, \Omega^{(1)}, \ldots, \Omega^{(r)}$ be the weight-increasing ordering of the equal weight partition of 2^N. Set $q_t = w(S)$ for any $S \in \Omega^{(t)}$, $t = 0, 1, 2, \ldots, r$. Then there is a finite number $r \leq 2^n - 1$ of marginal quotas q_t and corresponding intervals of stable power $(q_{t-1}, q_t]$ such that $W[N, q_t, \mathbf{w}] \subset W[N, q_{t-1}, \mathbf{w}]$.*

Proof Follows from the fact that card $(2^N) = 2^n$ and an increasing series of k real numbers $a_1, \ldots, a_k$ subdivides interval $(a_1, a_k]$ into $k - 1$ segments. An analysis of voting power as a function of the quota (given voting weights) can be substituted by an analysis of voting power in a finite number of marginal quotas. □

Proposition 4 *Let $[N, q, w]$ be a simple quota weighted committee and $(q_{t-1}, q_t]$ is the interval of stable power for quota q. Then for any $\gamma = 1 - q_t + \varepsilon$, where $\varepsilon \in (0, q_t - q_{t-1}]$ and for all $i \in N$*

$$card(C_i(N, q, \mathbf{w})) = card(C_i(N, \gamma, \mathbf{w}))$$

and

$$card(P_i(N, q, \mathbf{w})) = card(P_i(N, \gamma, \mathbf{w}))$$

Proof Let S be a winning coalition, k has the swing in S and $(q_{t-1}, q_t]$ is an interval of stable power for q. Then it is easy to show that $N \backslash S \cup k$ is a winning coalition, k has a swing in $N \backslash S \cup k$ and $(1 - q_t, 1 - q_{t-1}]$ is an interval of stable power for any quota $\gamma = 1 - q_t + \varepsilon$ $(0 < \varepsilon \leq q_t - q_{t-1})$. Let R be a winning coalition, j has a swing in R, and $(1 - q_t, 1 - q_{t-1}]$ is an interval of stable power for quota $\gamma = 1 - q_t + \varepsilon$ $(0 < \varepsilon \leq q_t - q_{t-1})$. Then $N \backslash R \cup j)$ is a winning coalition, j has a swing in N\R∪j and $(q_{t-1}, q_t]$ is an interval of stable power for any quota $q = q_{t-1} + \tau$ where $0 < \tau \leq q_t - q_{t-1}$. □

While in [N, q, **w**] the quota q means the total weight necessary to pass a proposal (and therefore we can call it a *winning quota*), the *blocking quota* means the total weight necessary to block a proposal. If q is a winning quota and $(q_{t-1}, q_t]$ is a quota interval of stable power for q, then any voting quota $1 - q_{t-1} + \varepsilon$ (where $0 < \varepsilon \leq q_t - q_{t-1}$), is a blocking quota. From Proposition 4 it follows that the blocking power of the committee members, measured by swing and pivot-based power indices, is equal to their voting power. It is easy to show that voting power and blocking power might not be the same for power indices based on membership in minimal winning coalitions (HP and DP power indices). Let r be the number of marginal quotas, then from Proposition 4 it follows that for power indices based on swings and pivots the number of majority power indices does not exceed *int* (*r/2*) + *1*.

Proposition 5 *Let* $q_1, q_2, \ldots, q_m$ *be the set of all majority marginal quotas in a simple weighted committee* [*N,* **w**], *and* $\boldsymbol{\pi}^k$ *be a vector of Shapley-Shubik relative power indices corresponding to a marginal quota* q_k, *then there exists a vector* $(\lambda_1, \lambda_2, \ldots, \lambda_r)$ *such that:*

$$\sum_{k=1}^{m} \lambda_k = 1, \lambda_k \geq 0, \sum_{k=1}^{m} \lambda_k \boldsymbol{\pi}^k = \mathbf{w}$$

The proof follows from Berg and Holler (1986). They provide the following property of simple weighted committees: Let [N, Q, **w**] be a finite family of simple quota weighted committees with the same weights **w** and a finite set of different relative quotas $Q = \{q_1, q_2, \ldots, q_m\}$. Let $\lambda(Q)$ be a probability distribution over Q where j_k is a probability with which a random mechanism selects the quota q_k and $\pi_{ik}(N, q_k, \mathbf{w})$ be SS relative power index in the committee $[N, q_k, \mathbf{w}]$ with a quota $q_k \in Q$, then

$$\bar{\pi}_i(N, Q, \mathbf{w}) = \sum_{k: q_k \in Q} \pi_{ik}(N, q_k, \mathbf{w}) \lambda_k$$

is an expected SS relative power of the member i in the randomized committee [N, $\boldsymbol{\lambda}$(Q), **w**]. For any vector of weights there exist a finite set Q of quotas q_k such that $0.5 < q_k \leq 1$, and a probability distribution $\boldsymbol{\lambda}$ such that

$$\bar{\pi}_i(N, Q, \mathbf{w}) = \sum_{k: q_k \in Q} \pi_{ik}(N, q_k, \mathbf{w}) \lambda_k = w_i$$

The randomized voting rule $\boldsymbol{\lambda}$(Q) leads to strict proportional expected SS power. Clearly, if there exists an exact quota q* such that $\pi_i(N, q*, \mathbf{w}) = w_i$, we can find it among finite number of marginal majority quotas.

In general, the number of majority power indices can be greater than the number of committee members, and the system

$$\sum_{k=1}^{r} \lambda_k = 1, \lambda_k \geq 0, \sum_{k=1}^{r} \lambda_k \boldsymbol{\pi}^k = \mathbf{w}$$

might not have the unique solution. To solve the system we can use the optimization problem:minimize

$$\sum_{i=1}^{n} abs\left(\sum_{k=1}^{r} \pi_i^k \lambda_k - w_i\right)$$

subject to

$$\sum_{k=1}^{r} \lambda_k = 1, \lambda_k \geq 0$$

that can be transformed into an equivalent linear programming problem (see Gale 1960):minimize

$$\sum_{i=1}^{n} y_i$$

subject to

$$\begin{aligned}
&\sum_{k=1}^{r} \pi_i^k \lambda_k - y_i \leq w_i && \mathit{for}\ i = 1, \ldots, n \\
&\sum_{k=1}^{r} \pi_i^k \lambda_k + y_i \geq w_i && \mathit{for}\, i = 1, \ldots n \\
&\sum_{k=1}^{r} \lambda_k = 1 && \\
&\lambda_k, y_i \geq 0 && \mathit{for}\, k = 1, \ldots, r, \quad i = 1, \ldots, n
\end{aligned}$$

This problem is easy to solve by standard linear programming simplex methods.

Although we can apply a randomized voting rule to any relative power index, based on pivots and swings, the problem is with the interpretation of what we get. The relative PB index has no probabilistic interpretation, so the randomized voting rule calculated for it by Proposition 5 does not provide the mathematical expectation of the number of swings, leading to a relative PB power equal to weights.

Moreover, one can hardly expect that randomized voting rules leading to the strict proportional expectation of power would be adopted by actors in real voting systems. However, the design of a "fair" voting system can be based on an approximation provided by the quota generating the minimal distance between vectors of power indices and weights, which is called an *optimal quota*.

The optimal quota was introduced by Słomczyński and Życzkowski (2006, 2007) and Turnvoec (2011) as a quota minimizing the sum of square residuals between the power indices and the voting weights by q ∈ (0.5, 1)

$$\sigma^2(q) = \sum_{i \in N} (\pi_i[N, q, \mathbf{w}] - w_i)^2$$

Słomczyński and Życzkowski introduced the optimal quota concept within the framework of the so-called Penrose voting system as a principle of fairness in the EU Council of Ministers voting. Here power is measured by the Penrose-Banzhaf power index. The system consists of two rules:

(a) The voting weight attributed to each member of the voting body of size n is proportional to the square root of the population he or she represents;
(b) The decision of the voting body is taken if the sum of the weights of members supporting it is not less than the optimal quota.

Looking for a quota providing a priori voting power "as close as possible" to the normalized voting weights, Słomczyński and Życzkowski (Turnovec 2011 in this volume) are minimizing the sum of square residuals between the power indices and voting power for q ∈ (0.5, 1]. Based on a simulation they propose heuristic approximations of the solution for the PB index:

$$\underline{q} = \frac{1}{2}\left(1 + \frac{1}{\sqrt{n}}\right) \leq q \leq \frac{1}{2}\left(1 + \sqrt{\sum_{i \in N} w_i^2}\right) = \overline{q}$$

Clearly $\underline{q} = \overline{q}$ if and only if all the weights are equal, but in this case any majority quota is optimal.

Definition 5 By index of the fairness of a voting rule q in [N, q, **w**] we call:

$$\phi(N, q, \mathbf{w}) = 1 - \sqrt{\frac{1}{2}\sum_i (\pi_i[N, q, \mathbf{w}] - w_i)^2}.$$

It is easy to see that $0 \leq \sqrt{\frac{1}{2}\sum_i (\pi_i[N, q, \mathbf{w}] - w_i)^2} \leq 1$ (zero in the case of the equality of weights and power, e.g., $w_1 = 1/2$, $w_2 = 1/2$, $\pi_1 = 1/2$, $\pi_2 = 1/2$, and 1 in the case of an extreme inequality of weights and power, e.g., $w_1 = 1$, $w_2 = 0$, $\pi_1 = 0$, $\pi_2 = 1$), hence $0 \leq \varphi(N, q, w) \leq 1$. We say that a voting rule q_1 is "at least as fair" as a voting rule q_2 if $\varphi(N, q_1, w) \geq \varphi(N, q_2, w)$.[5]

Looking for a "fair" voting rule we can maximize φ which is the same as to minimize σ^2(q). Using marginal quotas and intervals of stable power we do not need any simulation.

[5] The index of fairness follows the same logic as measures of deviation from proportionality used in political science to evaluate the difference between results of an election and the composition of an elected body—e.g., the measure given in Loosemore and Hanby (1971) is based on the absolute values of the deviation metric, or Gallagher (1991) using a square roots metric.

Proposition 6 *Let* $[N, q, w]$ *be a simple quota-weighted committee and* $\pi_i\,(N, q_t, w)$ *be relative power indices for marginal quotas* q_t, *and* q_t* *be the majority marginal quota minimizing*

$$\sum_{i \in N} \left(\pi_i\left(N, q_j, \mathbf{w}\right) - w_i\right)^2$$

($j = 1, 2, \ldots, r$, *r is the number of intervals of stable power such that* q_j *are marginal majority quotas*), *then the exact solution of Słomczyński and Życzkowski's optimal quota* (*SZ optimal quota*) *problem for a particular power index used is any* $\gamma \in (q_{t-1}^*, q_t^*]$ *from the quota interval of stable power for* q_t^*.

The proof follows from the finite number of quota intervals of stable power (Proposition 4). The quota q* provides the best approximation of strict proportional power that is related neither to a particular power measure nor to a specific principle of fairness.

4 Fair Quota in the Lower House of the Czech Parliament

To illustrate the concept of fair quota we use the structure of the recent term (2010–2014) of the Lower House of the Czech Parliament. The Lower House has 200 seats. Members of the Lower House are elected in 14 electoral districts from party lists by a proportional system with a 5 % threshold. Seats are allocated to the political parties that obtained not less than 5 % of total valid votes roughly proportionally to fractions of obtained votes (votes for parties not achieving the required threshold are redistributed among the successful parties roughly proportionally to the shares of obtained votes). Five political parties qualified to the Lower House in 2010: left centre Czech Social Democratic Party (Česká strana sociálně demokratická, ČSSD), right centre Civic Democratic Party (Občanská demokratická strana, ODS), right TOP09 (Tradice, Odpovědnost, Prosperita—Traditions, Responsibility, Prosperity 2009), left Communist Party of Bohemia and Moravia (Komunistická strana Čech a Moravy, KSČM) and supposedly centre (but not very clearly located on left–right political dimension) Public Issues (Věci veřejné, VV).

Table 1 provide results of the 2010 Czech parliamentary election. (By relative voting weights we mean fractions of seats of each political party, by relative electoral support fractions of votes for political parties that qualified to the Lower House, counted from votes that were considered in allocation of seats.) Three parties, ODS, TOP09 and VV, formed a right-centre government coalition with 118 seats in the Lower House.

We assume that all Lower House members of the same party vote together and all of them participate in each voting act. Two voting rules are used: simple majority (more than 100 votes) and qualified majority (at least 120 votes). There exist 16 possible winning coalitions for simple majority voting (12 of them are winning coalitions for qualified majority), 16 marginal majority quotas and

Table 1 Results of 2010 election to the lower house of the Czech parliament

	Seats	Votes in % of valid votes	Relative voting weight	Relative electoral support
ČSSD	56	22.08	0.28	0.273098
ODS	53	20.22	0.265	0.250093
TOP09	41	16.7	0.205	0.206555
KSČM	26	11.27	0.13	0.139394
VV	24	10.58	0.12	0.13086
$\sum$	200	80.85	1	1

Source http://www.volby.cz/pls/ps2010/ps?xjazyk=CZ

Table 2 Possible winning coalitions in the lower house of the Czech parliament (own calculations)

Parties of possible winning coalitions	Absolute marginal majority quota	Relative marginal majority quota	Intervals of stable power
ODS + KSČM + VV	103	0.515	(0.485, 0.515]
CSSD + KSČM + VV	106	0.53	(0.515, 0.53]
ČSSD + ODS	109	0.545	(0.53, 0.545]
ODS + TOP09 + VV	118	0.59	(0.545, 0.59]
ODS + TOP09 + KSČM	120	0.6	(0.59, 0.6]
ČSSD + TOP09 + VV	121	0.605	(0.6, 0.605]
ČSSD + TOP09 + KSČM	123	0.615	(0.605, 0.615]
ČSSD + ODS + VV	133	0.665	(0.615, 0.665]
ČSSD + ODS + KSCM	135	0.675	(0.665, 0.675]
ODS + TOP09 + KSČM + VV	144	0.72	(0.675, 0.72]
ČSSD + TOP09 + KSČM + VV	147	0.735	(0.72, 0.735]
ČSSD + ODS + TOP09	150	0.75	(0.735, 0.75]
ČSSD + ODS + KSČM + VV	159	0.795	(0.75, 0.795]
CSSD + ODS + TOP09 + VV	174	0.87	(0.795, 0.87]
ČSSD + ODS + TOP09 + KSČM	176	0.88	(0.87, 0.88]
ČSSD + ODS + TOP09 + KSČM + VV	200	1	(0.88, 1]

16 majority quota intervals of stable power (see Table 2). For the analysis of fair voting rule we applied the Shapley-Shubik power index and an Euclidean distance function. In Table 3 we provide the Shapley-Shubik power indices (distribution of relative voting power) for all of marginal majority quotas.

For any quota from each of the intervals of stable power the Shapley-Shubik of relative power is identical with the relative power in the corresponding marginal majority quota.

The fair relative majority quota in our case is q = 0.675 (with index of fairness equal to 0.95589), or any quota from interval of stable power (0.665, 0.675]. It means that minimal number of votes to approve a proposal is 135 (in contrast to 101 votes required by simple majority and 120 votes required by qualified majority). Voting rule defined by this quota maximizes the index of fairness (measured for Shapley-Shubik power index) and approximately equalizes the

Table 3 Shapley-Shubik power of political parties for majority marginal quotas (own calculations)

Party	Seats	Relative voting weight	SS power for q = 0.515	SS power for q = 0.53	SS power for q = 0.545	SS power for q = 0.59	SS power for q = 0.6	SS power for q = 0.605	SS power for q = 0.615	SS power for q = 0.665
ČSSD	56	0.28	0.3	0.35	0.3167	0.2667	0.3167	0.3667	0.3333	0.3
ODS	53	0.265	0.3	0.2667	0.3167	0.2667	0.2333	0.2	0.25	0.3
TOP09	41	0.205	0.1333	0.1833	0.2333	0.2667	0.2333	0.2	0.1667	0.1333
KSČM	26	0.13	0.1333	0.1	0.0667	0.1	0.15	0.1167	0.1667	0.1333
VV	24	0.12	0.1333	0.1	0.0667	0.1	0.0667	0.1167	0.0833	0.1333
$\sum$	200	1	0.9999	1	1.0001	1.0001	1	1.0001	1	0.9999
Index of fairness			0.94104	0.94223	0.92362	0.94859	0.94346	0.92269	0.93989	0.94104

Party	Seats	Relative voting weight	SS power for q = 0.675	SS power for q = 0.72	SS power for q = 0.735	SS power for q = 0.75	SS power for q = 0.795	SS power for q = 0.87	SS power for q = 0.88	SS power for q = 1
ČSSD	56	0.28	0.2667	0.2333	0.4333	0.3833	0.35	0.3	0.25	0.2
ODS	53	0.265	0.2667	0.2333	0.1833	0.3833	0.35	0.3	0.25	0.2
TOP09	41	0.205	0.1833	0.2333	0.1833	0.1333	0.1	0.3	0.25	0.2
KSČM	26	0.13	0.1833	0.15	0.1	0.05	0.1	0.05	0.25	0.2
VV	24	0.12	0.1	0.15	0.1	0.05	0.1	0.05	0	0.2
$\sum$	200	1	1	0.9999	0.9999	0.9999	1	1	1	1
Index of fairness			0.95589	0.94859	0.87361	0.85664	0.88943	0.89524	0.87361	0.89524

voting power (influence) of the members of the Lower House independently of their political affiliation.

5 Concluding Remarks

In simple quota weighted committees with a fixed number of members and voting weights there exists a finite number r of different quota intervals of stable power ($r \leq 2^n - 1$) generating a finite number of power indices vectors. For power indices with a voting power equal to blocking power the number of different power indices vectors corresponding to majority quotas is equal to int $(r/2) + 1$ at most.

If the fair distribution of voting weights is defined, then the fair distribution of voting power is achieved by the quota that maximizes the index of fairness (minimizes the distance between relative voting weights and relative voting power). The index of fairness is not a monotonic function of the quota.

The problem of optimal quota has an exact solution via the finite number of majority marginal quotas. Słomczyński and Życzkowski introduced an optimal quota concept within the framework of the so called Penrose voting system as a principle of fairness in the EU Council of Ministers voting and related it exclusively to the Penrose-Banzhaf power index and the square root rule. However, the fairness in voting systems and approximation of strict proportional power is not exclusively related to the Penrose square-root rule and the Penrose-Banzhaf definition of power, as it is usually done in discussions about EU voting rules. In this chapter it is treated in a more general setting as a property of any simple quota weighted committee and any well-defined power measure. Fairness and its approximation by optimal quota are not specific properties of the Penrose-Banzhaf power index and square root rule.

Acknowledgments This research was supported by the Czech Science Foundation, project No. 402/09/1066 “Political economy of voting behavior, rational voters’ theory and models of strategic voting” and by the Max Planck Institute of Economics in Jena. The author would like to thank Manfred J. Holler, Andreas Nohn and an anonymous referee for valuable comments to an earlier version of the chapter.

References

Baldwin, R., & Widgrén, M. (2004). Winners and losers under various dual majority rules in the EU’s council of ministers. In M. Wiberg (Ed.), *Reasoned choices* (pp. 42–89). Turku: The Finnish Political Science Association.

Banzhaf, J. F. (1965). Weighted voting doesn’t work: A mathematical analysis. *Rutgers Law Review, 19*, 317–343.

Berg, S., & Holler, M. J. (1986). Randomized decision rules in voting games: A model for strict proportional power. *Quality and Quantity, 20*, 419–429.

Brueckner, M. (2001). Expected probabilistic characterization of power indices. In M.J. Holler & G. Owen (Eds.), *Power indices and coalition formation*, Boston-Dordrecht-London: Kluwer.

Deegan, J., & Packel, E. W. (1978). A new index of power for simple n-person games. *International Journal of Game Theory, 7*, 113–123.

Felsenthal, D.S., & Machover, M. (1998). *The Measurement of Voting Power, Theory and Practice*, Cheltenham: Edward Elgar.

Felsenthal, D.S., & Machover, M. (2004), *Analysis of QM rule adopted by the EU intergovernmental conference, Brussels,* 18 *June* 2004, Research report online, London: LSE. [http://eprints.lse.ac.uk/2531].

Felsenthal, D.S., & Machover, M. (2007). *Analysis of QM rule adopted by the council of the European Union, Brussels,* 23 *June* 2007, Research report online, London: LSE. [http://eprints.lse.ac.uk/archive0000431].

Felsenthal, D. S., Machover, M., & Zwicker, W. (1998). The bicameral postulates and indices of a priori voting power. *Theory and Decision, 44*, 83–116.

Gale, D. (1960). *The theory of linear economic models*. New York: McGraw-Hill.

Gallagher, M. (1991). Proportionality, disproportionality and electoral systems. *Electoral Studies, 10*, 33–51.

Holler, M. J. (1982a). An introduction to the analysis of power, voting, and voting power. In M. J. Holler (Ed.), *Power, voting and voting power* (pp. 15–30). Würzburg, Wien: Physica Verlag.

Holler, M. J. (1982b). Forming coalitions and measuring voting power. *Political Studies, 30*, 262–271.

Holler, M. J. (1985). Strict proportional power in voting bodies. *Theory and Decision, 19*, 249–258.

Holler, M. J. (1987). Paradox proof decision rules in weighted voting. In M. J. Holler (Ed.), *The logic of multiparty systems* (pp. 425–436). Dordrecht: Martinus Nijhoff Publishers.

Holler, M. J., & Packel, E. (1983). Power, luck and the right index. *Zeitschrift für Nationalőkonomie (Journal of Economics), 43*, 21–29.

Hosli, M. O. (2008). Council decision rules and european union constitutional design. *AUCO Czech Economic Review, 2*(1), 76–96.

Johnston, R. J. (1978). On the measurement of power: Some reaction to laver. *Environment and Planning, A10*, 907–914.

Laruelle, A., & Widgrén, M. (1998). Is the allocation of voting power among EU states fair? *Public Choice, 94*, 317–339.

Leech, D., & Aziz, H. (2008). T*he double majority voting rule of the EU reform treaty as a democratic ideal for an enlarging union: An appraisal using voting power analysis,* Presented at the 2008 EPCS meeting in Jena.

Loosemore, J., & Hanby, J. V. (1971). The theoretical limits of maximum some analytical expression for electoral systems. *British Journal of Political Science, 1*, 467–477.

Nurmi, H. (1982). The problem of the right distribution of voting power. In M. J. Holler (Ed.), *Power, voting and voting power* (pp. 203–212). Würzburg, Wien: Physica Verlag.

Nurmi, H. (1997a). The representation of voter groups in the European parliament. A Penrose-Banzhaf index analysis. *Electoral Studies, 16*, 317–322.

Nurmi, H. (1997b). On power indices and minimal winning coalitions. *Control and Cybernetics, 26*(4), 609–611.

Owen, G. (1972). Multilinear extensions of games. *Management Science, 18*, 64–79.

Penrose, L. S. (1946). The elementary statistics of majority voting. *Journal of the Royal Statistical Society, 109*, 53–57.

Shapley, L. S., & Shubik, M. (1954). A method for evaluation the distribution of power in a committee system. *American Political Science Review, 48*, 787–792.

Słomczyński, W., & Życzkowski, K. (2006). Penrose voting system and optimal quota. *Acta Physica Polonica, B37*(11), 3133–3143.

Słomczyński, W., & Życzkowski, K. (2007). From a toy model to double square root system. *Homo Oeconomicus, 24*(3–4), 381–400.

Turnovec, F. (1997). Power, power indices and intuition. *Control and Cybernetics, 26*(4), 613–615.

Turnovec, F. (2007). New measure of voting power. *AUCO Czech Economic Review, 1*(1), 4–14.

Turnovec, F. (2008a). National, political and institutional influence in European Union decision making. *AUCO Czech Economic Review, 2*(2), 154–173.

Turnovec, F. (2008b). Efficiency of fairness in voting systems. *Homo Oeconomicus, 25*(3–4), 335–348.

Turnovec, F. (2009). Fairness and squareness: Fair decision making rules in the EU council. *Badania Operacyjne i Decyzje, 4*, 109–124.

Turnovec, F. (2011). Fair voting rules in committees, strict proportional power and optimal quota. *Homo Oeconomicus, 27*(4), 463–479.

Turnovec, F., Mercik, J. W., & Mazurkiewicz, M. (2008). Power indices methodology: Decisiveness, pivots and swings. In M. Braham & F. Steffen (Eds.), *Power, freedom and voting, essays in honor of Manfred J. Holler* (pp. 23–37). Berlin, Heidelberg: Springer.

Widgrén, M. (2001). On the probabilistic relationship between the public good index and the normalized Banzhaf index. In M. J. Holler & G. Owen (Eds.), *Power indices and coalition formation*. Boston-Dordrecht-London: Kluwer.

On Penrose's Square-Root Law and Beyond

Werner Kirsch

1 Introduction

All modern democracies rely on the idea of representation. A certain body of representatives, a parliament for example, makes decisions on behalf of the voters. In most parliaments each of its members represents roughly the same number of people, namely the voters in his or her constituency.

There are other bodies in which the members represent different numbers of voters. A prominent example is the Council of the European Union. Here ministers of the member states represent the population of their respective country. The number of people represented in the different states differs from about 400,000 for Malta to more than 82 million for Germany. Due to this fact the members of the Council have a certain number of votes depending on the size of the country they represent, e.g. 3 votes for Malta, 29 votes for Germany. The votes of a country cannot be split, but have to be cast as a block.[1]

Similar voting systems occur in various other systems, for example in the Bundesrat, Germany's state chamber of parliament and in the electoral college in the USA.[2]

Let us call such a system in which the members represent subsystems (states) of different size a *heterogeneous voting system*. In the following we will call the

This chapter was originally published in Homo Oeconomicus 24 (3/4), 2007.

[1] The current voting system in the Council is based on the treaty of Nice. It has additional components to the procedure described above, which are irrelevant in the present context. For a description of this voting system and further references see e.g. Kirsch (2012).

[2] The electoral college is not exactly a heterogeneous voting system in the sense defined below, but it is very close to it.

W. Kirsch (✉)
Fakultät für Mathematik und Informatik, FernUniversität in Hagen, Hagen, Germany
e-mail: werner.kirsch@fernuni-hagen.de

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_20,

assembly of representatives in a heterogeneous voting system the *council*, the sets of voters represented by the council members the *states*.

It is quite clear, that in a heterogeneous voting system a bigger state (by population) should have at least as many votes in the council as a smaller state. It may already be debatable whether the bigger states should have *strictly* more votes than the smaller states (cf. the Senate in the US constitution). And if yes, how much more votes the bigger state should get?

In this note we address the question: ‘What is a fair distribution of power in a heterogeneous voting system?’

There exist various answers to this question, depending on the interpretation of the words ‘fair’ and ‘power’.

The usual and quite reasonable way to formulate the question in an exact way is to use the concept of power indices. One calls a heterogeneous voting system fair if all voters in the member states have the same influence on decisions of the council. By ‘same influence’ we mean that the power index of each voter is the same regardless of her or his home state. If we choose then Banzhaf power index to measure the influence of a voter we obtain the celebrated Penrose’s square-root law (see e.g. Felsenthal and Machover 1998).

The square-root law states that the distribution of power in a heterogeneous voting system is fair if the power (index) of each council member i is proportional to $\sqrt{N_i}$, where N_i is the population of the state which i represents.

In their book Felsenthal and Machover (1998) formulate a second square-root law. There they base the notion of ‘fairness’ on the concept of *majority deficit*.

The majority deficit is zero if the voters favoring the decision of the council are the *majority*. If the voters favoring the decision of the council are the *minority* then the majority deficit is the margin between the number of voters objecting to the decision and those agreeing with it (see Definition 3.3.16 in Felsenthal and Machover (1998)).

The notion of fairness we propose in this chapter is closely related to the concept of majority deficit. We will call a decision of the council *in agreement with the popular vote* if the percentage of voters agreeing with a proposal (popular vote) is as close as possible to the percentage of council votes in favor of the proposal. (We will make this notion precise in the next section.)

For both concepts we have to average over the possible voting configurations. This is usually done by assuming that voters vote independently of each other. The main purpose of this note is to investigate some (we believe reasonable) models where voters do not vote independently.

We will discuss two voting models with voting behavior which is *not* independent. The first model considers societies which have some kind of ‘collective bias’ (or ‘common belief’). A typical situation of this kind is a strong religious group (or church) influencing the voting behavior of the voters. This model is discussed in detail in Sect. 3.

In the other model voters tend to vote the same way ‘the majority does’. This is a situation where voters do not want to be different from others. We call this the *mean field model* referring to an analogous model from statistical physics. See Sect. 5 for this model.

In fact, both models can be interpreted in terms of statistical physics. Statistical physics considers (among many other things) magnetic systems. The elementary magnet, called a spin, has two possible states which are '$+1$' or '-1' (spin up, spin down). This models voting 'yes' or 'no' in a voting system. Physicists consider different kinds of interactions between the single spins, one given through an exterior magnetic-field—corresponding to a society with 'a collective bias'—or through the tendency of the spins to align—corresponding to the second voting model. We discuss the analogy of voting models with spin systems in Sect. 4.

Our investigations of voting models with statistical dependence is much inspired by the chapter Laruelle and Valenciano (2005). The first model is also based on the work by Straffin (1982).

It does not come as a surprise that we obtain a square-root law for a model with independent voters, just as in the case considered by Felsenthal and Machover (1998).

For the mean field model we still get a square-root law for the best possible representation in the council *as long as* the mutual interaction between voters is not too strong.

However as the coupling between voters exceeds a certain threshold, the fairest representation in the council is no longer given by votes proportional to $\sqrt{N_i}$ but rather by votes proportional to N_i. This is a typical example of a phase transition.

In the model of collective bias the fair representation weight depends on the strength of the collective bias for large populations. If this strength is independent of the population size fair representation is almost always given by voting weights proportional to N_i, the square-root law occurring only in marginal cases. However, if the collective bias decreases with increasing population one can get any power law behavior N_i^{α} for the optimal weight as long as $\frac{1}{2} \leq \alpha \leq 1$. In fact, statistical investigations on real life data suggest that this might happen (see Gelman et al. 2004).

We leave the mathematical proofs of our results for the appendices (Sects. 8–10).

2 The General Model

We consider N voters, denoted by $1, 2, \ldots, N$. Each of them may vote 'yes' or 'no'; abstentions are not allowed. The vote of the voter i is denoted by X_i.

The possible voting results are $X_i = +1$ representing 'yes' and $X_i = -1$ for 'no'. We consider the quantity X_i as random, more precisely there is a probability measure $\mathbb{P}$ on the space $\{-1, 1\}^N$ of possible voting results. We will call the measure $\mathbb{P}$ a *voting measure* in the following. $\mathbb{P}$ and its properties will be specified later. The conventional assumption on $\mathbb{P}$ is that the random quantities X_i are independent from each other, but we are *not* making this assumption here.

Our interpretation of this model is as follows. The voters react on a proposal in a rational way, that is to say: A voter does *not* roll a dice to determine his or her voting behavior but he or she votes for or against a given proposal according to his/her personal belief, knowledge, experience etc. It is rather the proposal which is the source of randomness in this system. We imagine the voting system is fed with propositions in a completely random way. This could be either a real source of proposals or just a Gedankenexperiment to measure the behavior of the voting system.

The rationality of the voters implies that a voter who casts a 'yes' on a certain proposition will necessarily vote 'no' on the diametrically opposed proposition. Since we assume that the proposals are completely random any proposal and its antithetic proposal must have the same probability. This implies

$$\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}. \tag{2.1}$$

More generally, we conclude that

$$\mathbb{P}(X_{i_1} = \xi_1, \ldots, X_{i_r} = \xi_r) = \mathbb{P}(X_{i_1} = -\xi_1, \ldots, X_{i_r} = -\xi_r) \tag{2.2}$$

for any set $i_1 \ldots, i_r$ of voters and any $\xi_1, \ldots \xi_r \in \{-1, 1\}$.

We call the property (2.2) the *symmetry* of the voting system. Any measure $\mathbb{P}$ satisfying (2.2) is called a *voting measure*.

The symmetry assumption (2.2) does *not* fix the probability measure $\mathbb{P}$. Only if we assume in addition that the X_i are statistically independent we can conclude from (2.2) that

$$\mathbb{P}(X_{i_1} = \xi_1, \ldots, X_{i_r} = \xi_r) = (\frac{1}{2})^r. \tag{2.3}$$

So far, we have not specified any decision rule for the voting system. The above probabilistic setup is completely independent from the voting rule, a fact which was emphasized in the work Laruelle and Valenciano (2005).

A *simple majority rule* for $X_1, \ldots, X_N$ is given by the decision rule: Accept a proposal if $\sum_{j=1}^{N} X_j > 0$ and reject it otherwise.

By a *qualified majority rule* we mean that at least a percentage q (called the *quota*) of votes is required for the acceptance of a proposal. In term of the X_j this means:

$$\sum_{j=1}^{N} X_j \geq (2q - 1)N. \tag{2.4}$$

Indeed, it is not hard to see that the number of affirmative votes is given by

$$\frac{1}{2}\left(\sum_{j=1}^{N} X_j + N\right).$$

From this the assertion (2.4) follows.

In particular, the simple majority rule is obtained form (2.4) by choosing q slightly bigger than $\frac{1}{2}$.

The sum $\sum_{j=1}^{N} X_j$ gives the difference between the number of 'yes'-votes and the number of 'no'-votes. We call the quantity

$$M(X) := \left| \sum_{j=1}^{N} X_j \right| \tag{2.5}$$

the margin of the voting outcome $X = (X_1, \ldots, X_N)$. It measures the size of the majority with which the proposal is either accepted or rejected in simple majority voting.

In qualified majority voting with quota q the corresponding quantity is the *q-margin* $M_q(X)$ given by:

$$M_q(X) := \left| \sum_{j=1}^{N} X_j - (2q-1)N \right|. \tag{2.6}$$

Now, we turn to voting in the council. We consider M states, the state number ν having N_ν voters. Consequently the total number of voters is $N = \sum N_\nu$. The vote of the voter i in state ν is denoted by $X_{\nu i}$, $\nu = 1, \ldots, M$ and $i = 1, \ldots, N_\nu$.[3]

We suppose that each state government knows the opinion of (the majority of) the voters in that state and acts accordingly.[4] That is to say: If the majority of people in state ν supports a proposal, i.e. if

$$\sum_{i=1}^{N_\nu} X_{\nu i} > 0 \tag{2.7}$$

then the representative of state ν will vote 'yes' in the council otherwise he or she will vote 'no'. If we set $\chi(x) = 1$ for $x > 0$, $\chi(x) = -1$ for $x \leq 0$ the representative of state ν will vote

$$\xi_\nu = \chi\left(\sum_{i=1}^{N_\nu} X_{\nu i} \right) \tag{2.8}$$

in the council. If the state ν has got a weight w_ν in the council the result of voting in the council is given by:

[3] We label the states using Greek characters and the voters within a state by Roman characters.

[4] Although this is the central idea of representative democracy this idealization may be a little naive in practice.

$$\sum_{\nu=1}^{M} w_\nu\, \xi_\nu = \sum_{\nu=1}^{M} w_\nu\, \chi\Biggl(\sum_{i=1}^{N_\nu} X_{\nu i}\Biggr). \tag{2.9}$$

Thus, the council's decision is affirmative if $\sum_{\nu=1}^{M} w_\nu\, \xi_\nu$ is positive, provided the council votes according to simple majority rule.

The result of a popular vote in all countries $\nu = 1,\ldots,N$ is

$$P = \sum_{\nu=1}^{M}\sum_{i=1}^{N_\nu} X_{\nu i}. \tag{2.10}$$

We will call voting weights w_ν for the council *fair* or *optimal*, if the council's vote is as close as possible to the public vote. To make this precise let us define

$$C = \sum_{\nu=1}^{M} w_\nu\, \chi\Biggl(\sum_{i=1}^{N_\nu} X_{\nu i}\Biggr) \tag{2.11}$$

the result of the voting in the council. Both P and C are random quantities which depend on the random variables $X_{\nu i}$. So, we may consider the mean square distance Δ between P and C, i.e. denoting the expectation over the random quantities by $\mathbb{E}$, we have

$$\Delta = \mathbb{E}\Bigl((P-C)^2\Bigr) \tag{2.12}$$

$$= \mathbb{E}\Biggl(\Biggl\{\sum_{\nu=1}^{M}\sum_{i=1}^{N_\nu} X_{\nu i} - \sum_{\nu=1}^{M} w_\nu \chi\bigl(\sum_{i=1}^{N_\nu} X_{\nu i}\bigr)\Biggr\}^2\Biggr). \tag{2.13}$$

In a democratic system the decision of the council should be as close as possible to the popular vote, hence we call a system of weights *fair* or *optimal* if $\Delta = \Delta(w_1,\ldots,w_M)$ is minimal among all possible values of w_ν.

In the following we suppose that the random variables $X_{\nu i}$ and $X_{\mu j}$ are independent for $\nu \neq \mu$. This means that voters in different states are not correlated. We do not assume at the moment that two voters from the same state vote independently of each other.

We have the following result:

Theorem 2.1. *Fair voting in the council is obtained for the values*

$$\begin{aligned} w_\nu &= \mathbb{E}\Biggl(\Biggl|\sum_{i=1}^{N_\nu} X_{\nu i}\Biggr|\Biggr) \\ &= \mathbb{E}\Bigl(M(X_\nu)\Bigr). \end{aligned}$$

This result can be viewed as an extension of Penrose's square-root law to the situation of correlated voters. We will see below that it gives $w_v \sim \sqrt{N_v}$ for independent voters.

Theorem 2.1 has a very easy—we hope convincing—interpretation: w_v is the expected margin of the voting result in state v. In other words, it gives the expected number of people in state v that agree with the voting of v in their council minus those that disagree, i.e. the net number of voters which the council member of v actually represents.

If we choose a multiple $c\,w_1, \ldots, c\,w_{N_v}$ $(c > 0)$ of the weights $w_1, \ldots w_{N_v}$ we obtain the same voting system as the one defined by $w_1, \ldots, w_n$. In this sense the weights w_v of Theorem 2.1 are not unique, but the voting system is.

We will prove Theorem 2.1 in Sect. 8. We remark that the proof requires the symmetry assumption (2.2) and the independence of voters from *different* states.

The next step is to compute the expected margin $\mathbb{E}(M(X_v))$, at least asymptotically for large number of voters N_v. This quantity depends on the correlation structure between the voters in state v. As we will see, different correlations between voters give very different results for $\mathbb{E}(M(X_v))$ and hence for the optimal weight w_v.

We begin with the classical case of independent voters.

Theorem 2.2. *If the voters in state v cast their votes independently of each other then*

$$\mathbb{E}\left(\left|\sum_{i=1}^{N_v} X_{vi}\right|\right) \sim c\,\sqrt{N_v} \tag{2.14}$$

for large N_v.

Thus, we recover the square-root law as we expected. (For the square-root law see Felsenthal and Machover (1998).) In terms of power indices the independence assumption is associated to the Banzhaf power index. Therefore, it is not surprising that also the Banzhaf index leads to a square-root rule.

It is questionable (as we know from the work of Gelman et al. (2004)) whether the independent voters model is valid in many real-life voting systems. This is one of the reasons to extend the model as we do in the present chapter.

3 The 'Collective Bias' Model

In this section we define and investigate a model we dub the 'collective bias model'. In this model there exists a kind of common opinion in the society from which the individual voters may deviate but to which they agree in the mean. Such a 'common opinion' or 'collective bias' may have very different reasons: There may be a system of common values or common beliefs in the country under consideration, there may be an influential religious group or political ideology, there could be a strong tradition or simply a common interest based on economical needs. A 'collective bias' may also originate in a single person's influence on the media or in the pressure put onto voters by some powerful group. The obviously important differences in the origin of the common opinion are not reflected by the model as the purely technical outcome does not depend on it.

To model the collective bias we introduce a random variable Z, the collective bias variable, which takes values between -1 and $+1$. If $Z > 0$ the collective bias is in favor of the proposition under consideration. The closer the value of Z to 1, the higher the expected percentage of voters in favor of the proposition. In particular, if $Z = 1$ *all* voters will vote 'Yes', while $Z = 0$ means the collective bias is neutral towards the proposal and the voters vote independent of each other and with probability one half for (or against) the proposal. In general, if the collective bias variable Z has the value ζ the probability that the voter i votes 'Yes' is $p_\zeta = \frac{1}{2}\,(1+\zeta)$, the probability for a 'No' is consequently $1 - p_\zeta = \frac{1}{2}\,(1-\zeta)$. The probability p_ζ is chosen such that the expectation value of X_i is ζ, the value of the 'collective bias' variable Z. Thus ζ equals the expected fraction of voters supporting the proposal.

We remark that Z is a random variable, which means it depends on the proposal under consideration. This models the fact that there may be a strong common belief on certain issues while there is no or merely a weak common opinion on others. For example, in a country with a strong influence of the catholic church there may be a strong common view about abortion among voters, but, perhaps, not about speed limits on highways or on the details of taxation.

Once the value ζ of Z is chosen the voters vote independently of each other but with a probability for 'Yes' and 'No' which depends on ζ. The voting results $(X_1, \ldots, X_N)$ are correlated through (and only through) the collective bias Z.

In the following we describe the 'collective bias model' in a formal way. We introduce a random variable Z (the 'collective bias') with values in the interval $[-1, 1]$ and with a probability distribution μ, which we call the 'collective bias measure'. $\mu([a, b])$ is the probability that Z takes a value in $[a, b]$. For a given $\zeta \in [-1, 1]$ we denote by P_ζ the probability measure on $\{-1, 1\}$ with:

$$\begin{aligned} & P_\zeta(X_i = 1) = p_\zeta = \frac{1}{2}(1+\zeta) \\ \text{and} \quad & P_\zeta(X_i = -1) = 1 - p_\zeta = \frac{1}{2}(1-\zeta) \end{aligned} \tag{3.1}$$

We set $p_\zeta = \frac{1}{2}\,(1+\zeta)$. p_ζ is chosen such that we have $E_\zeta(X_i) = \zeta$ where E_ζ denotes the expectation value with respect to P_ζ.

Now, we define the voting measure $\mathbb{P}_\mu$ with respect to the 'collective bias measure' μ. The *conditional probability* with respect to $\mathbb{P}_\mu$ *given* $Z = \zeta$ is obtained from:

$$\mathbb{P}_\mu(X_1 = \xi_1, \ldots, X_N = \xi_N \,|\, Z = \zeta) = \prod_{i=1}^{N} P_{p_\zeta}(X_i = \xi_i) \tag{3.2}$$

Thus, given the value ζ of the 'collective bias' variable Z, the voters vote independently of each other with expected outcome equal to ζ. As a consequence of (3.2), the measure $\mathbb{P}_\mu$ is given by integrating over ζ, hence:

$$\mathbb{P}_\mu(X_1 = \xi_1, \ldots, X_N = \xi_N) = \int \prod_{i=1}^{N} P_{p_\zeta}(X_i = \xi_i) d\mu(\zeta). \tag{3.3}$$

To ensure that the probability $\mathbb{P}_\mu$ satisfies the symmetry condition (2.2) we have to require that

$$\mathbb{P}_\mu(Z \in [a, b]) = \mathbb{P}_\mu(Z \in [-b, -a])$$

i.e.

$$\mu([a, b]) = \mu([-b, -a]) \tag{3.4}$$

The probability measure $\mathbb{P}_\mu$ defines a whole class of examples, each (symmetric) probability measure μ on $[-1, 1]$ defines its unique $\mathbb{P}_\mu$. For example, if we choose $\mu = \delta_0$, i.e. $\mu([a, b]) = 1$ if $a \le 0 \le b$ and $= 0$ otherwise, we obtain independent random variables X_i as discussed in the final part of Sect. 2. Indeed, $\mu = \Delta_0$ means that $Z = 0$, consequently (3.3) defines independent random variables. Observe, that this is the only measure for which Z assumes a fixed value, since the collective bias measure μ has to be symmetric (3.4).

Another interesting example is the case when μ is the uniform distribution on $[-1, 1]$, meaning that each value in the interval $[-1, 1]$ is equally likely. This case was considered by Straffin (1982). He observed that this model is intimately connected with the Shapley-Shubik power index. We will comment on this interesting connection and on Straffin's calculation in an appendix (Sect. 7). To apply the 'collective bias' model to a given heterogeneous voting model we have to specify the measure μ, of course. In fact, this measure may change from state to state. In particular, one may argue that larger states tend to have a less homogeneous population and hence, for example, the influence of a specific religious or political group will be smaller. As an example to this phenomenon, we will later discuss a model modifying Straffin's example where $\mu(dz) = \frac{1}{2}\chi_{[-1,1]}(z)\, dz$ (uniform distribution in $[-1, 1]$) to a measure where μ_N depends on the population N, namely

$$\mu_N(dz) = \frac{1}{2a_N} \chi_{[-a_N, a_N]}(z) dz \tag{3.5}$$

with parameters $0 < a_N \le 1$. In particular, if we have $a_N \to 0$ as $N \to \infty$, the parameter a_N reflects the tendency of a common belief to decrease with a growing population.

Except for the trivial case $\mu = \delta_0$ the random variables X_i are never independent under $\mathbb{P}_\mu$. This can be seen from the covariance

$$\langle X_i, X_j \rangle_\mu := \mathbb{E}_\mu(X_i X_j) - \mathbb{E}_\mu(X_i)\mathbb{E}_\mu(X_j). \tag{3.6}$$

In (3.6) as well as in the following $\mathbb{E}_\mu$ denotes expectation with respect to $\mathbb{P}_\mu$. In fact, the random variables X_i are always positively correlated:

Theorem 3.1. *For $i \neq j$ we have*

$$\langle X_i, X_j \rangle_\mu = \int \zeta^2 \, d\mu(\zeta). \tag{3.7}$$

The quantity $\int \zeta^2 \, d\mu(\zeta)$ is called the second moment of the measure μ. Since the first moment $\int \zeta \, d\mu(\zeta)$ vanishes due to (3.4) the second moment equals the variance of μ. Observe that $\int \zeta^2 d\mu(\zeta) = 0$ implies $\mu = \delta_0$. For independent random variables $\langle X_i, X_j \rangle_\mu = 0$, so (3.7) implies that X_i, X_j depend on each other unless $\mu = \delta_0$.

To investigate the impact of the collective bias measure μ on the ideal weight in a heterogeneous voting model we have to compute the quantity

$$\mathbb{E}_\mu(|\sum_{i=1}^{N} X_i|) \tag{3.8}$$

for a measure μ and population N (at least for large N). This is done with the help of the following Theorem:

Theorem 3.2 *We have*:

$$\Big| \, \mathbb{E}_\mu(\frac{1}{N} \, |\sum_{i=1}^{N} X_i \,|) - \int \, |\zeta| \, d\mu(\zeta) \Big| \leq \frac{1}{\sqrt{N}} \tag{3.9}$$

Let us define $\overline{\mu} = \int |\zeta| \, d\mu(\zeta)$. If we choose $\mu \neq \delta_0$ independent of the (population of the) state Theorem 3.2 implies that the optimal weight in the council is *proportional* to N (rather than $\sqrt{N}$). This is true in particular for the original Straffin model (Straffin 1982) where $\mu_n \equiv \frac{1}{2} \chi_{[-1,1]}(z) \, dz$ which corresponds to the Shapley-Shubik power index (see Sect. 7). We have:

Theorem 3.3 *If the collective bias measure $\mu \neq \delta_0$ is independent of N then the optimal weight in the council is given by*:

$$w_N = \mathbb{E}_\mu(\sum_{i=1}^{N} \, |X_i|) \sim \overline{\mu} N \tag{3.10}$$

If $\mu = \mu_N$ depends on the population then

$$\mathbb{E}_{\mu_N}(|\sum_{i=1}^{N} X_i|) \sim N\overline{\mu}_N$$

as long as $\overline{\mu}_N \geq \frac{1}{N^{1/2-\varepsilon}}$ for some $\varepsilon > 0$. However, if $\overline{\mu}_N \leq \frac{1}{N^{1/2+\varepsilon}}$, then

$$E_{\mu_N}(|\sum_{i=1}^{N} X_i|) \sim \sqrt{N}.$$

Hence, in this case we rediscover a square-root law.

We summarize:

Theorem 3.4 *Let us suppose that a state with a population of size N is characterized by a collective bias measure μ_N, then*:

(1) *If*

$$\overline{\mu}_N = \int |\zeta|\, d\mu_N(\zeta) \geq C\, \frac{1}{N^{1/2-\varepsilon}} \tag{3.11}$$

for some $\varepsilon > 0$ *and for all large N then the optimal weight w_N is given by*:

$$w_N = \mathbb{E}_{\mu}(|\sum_{i=1}^{N} X_i|) \sim N\,\overline{\mu}_N. \tag{3.12}$$

(2) *If*

$$\overline{\mu}_N = \int |\zeta| d\mu_N(\zeta) \leq C\, \frac{1}{N^{1/2+\varepsilon}} \tag{3.13}$$

then for large N the optimal weight w_N is given by:

$$w_N = \mathbb{E}_{\mu}(|\sum_{i=1}^{N} X_i|) \sim \sqrt{N}. \tag{3.14}$$

Example: In our Straffin-type example (3.5) we choose:

$$\mu_N(dz) = \frac{1}{2a_N}\chi_{[-a_N,a_N]}(z)dz, \tag{3.15}$$

then:

$$\overline{\mu}_N = \frac{1}{2}\, a_N. \tag{3.16}$$

Let us assume $a_N \sim C\,N^{-\alpha}$ for $0 \leq \alpha \leq 1$. Then, if $\alpha > \frac{1}{2}$ we have $w_N \sim \sqrt{N}$ and if $\alpha < \frac{1}{2}$ we obtain $w_N \sim C\,N^{1-\alpha}$.

Remarks 3.5

(1) Our result shows that in all cases the optimal weight w_N satisfies $C\sqrt{N} \leq w_N \leq N$. It is a matter of empirical studies to determine which measure μ_N is appropriate to the given voting system. Any of the empirical results of Gelman et al. (2004) can be modeled by an appropriate choice of μ_N.
(2) It is only $\overline{\mu}_N$ that enters the formulae (3.12) and (3.14), no other information about μ_N is relevant. The quantities $\overline{\mu}_N$ can be estimated using Theorem 3.2. In fact, more is true by the following result.

Theorem 3.6 *Let P_N be the distribution of $\frac{1}{N}\sum_{i=1}^{N} X_i$ under the measure $\mathbb{P}_{\mu_N}$ then the sequence of measures $P_N - \mu_N$ converges weakly to* 0.

Theorem 3.6 tells us that the collective bias measure can be recovered from voting results. Let us denote by S a voting result, i.e. $S = \frac{1}{N}\sum_{i=1}^{N} X_i$. In other words, $\frac{1}{2}(S+1)$ is the fraction of affirmative votes. Theorem 3.6 tells us that the probability distribution of S approximates the measure μ_N for large N. On the other hand the distribution of S can be estimated from independent voting samples (defining the empirical distribution).

Note that the empirical distribution of voting results $\frac{1}{N}\sum_{i=1}^{N} X_i$ is the quantity considered in Gelman et al. (2004). Theorem 3.6 tells us that the distribution of the voting results for large number N of voters is approximately equal to the distribution μ_N. In particular, in the case of independent voters the voting result is always extremely tight while for Straffin's example any voting result has the same probability, i.e. it is equally likely that a proposal gets 99 or 53 % of the votes. The general 'collective bias model' defined above is an extension both of the independent voting model and of Straffin's model. This general model can be fit to any distribution of voting results.

4 Voting Models as Spin Systems

Spin systems are a central topic in statistical physics. They model magnetic phenomena. The spin variables, usually denoted by σ_i, may take values in the set $\{-1, +1\}$ with $+1$ and -1 meaning 'spin up' and 'spin down' respectively. The spin variables model the elementary magnets of a material (say the electrons or nuclei in a solid). The index i runs over an index set I which represents the set of elementary magnets. We may (and will) take $I = \{1, 2, \ldots, N\}$ in the following.

A *spin configuration* is a sequence $\{\sigma_i\}_{i\in\{1,\ldots,N\}} \in \{-1, +1\}^N$. A configuration of spins $\{\sigma_i\}_{i\in I}$ has a certain *energy* which depends on the way the spins interact with each other and (possibly) an exterior magnetic field. The energy, a function of the spin configuration, is usually denoted by $\mathcal{E} = \mathcal{E}(\{\sigma_i\}_{i\in I})$. Spin systems prefer configurations with small energy. For example in so called ferromagnetic systems, magnetic materials we encounter in every day's life, the energy of spins pointing in the same direction is smaller than the one for antiparallel spins, hence there is a tendency that spins line up, a fact that leads to the existence of magnetic materials.

The temperature T measures the strength of fluctuations in a spin system. If the temperature T of a system is zero, there are no fluctuations and the spins will stay in the configuration(s) with the smallest energy. However, if the temperature is positive, there is a certain probability that the spin configuration deviates from the one with the smallest energy. The probability to find a spin system at temperature $T > 0$ in a configuration $\{\sigma_i\}$ is given by:

$$p(\{\sigma_i\}) = \mathcal{Z}^{-1} e^{-\frac{1}{T}\mathcal{E}(\{\sigma_i\})} \tag{4.1}$$

The quantity $\mathcal{Z}$ is merely a normalization constant, to ensure that the right hand side of (4.1) defines a probability (i.e. gives total probability equal to one). Consequently:

$$\mathcal{Z} = \sum_{\{\sigma_i\}\in\{-1,1\}^N} e^{-\frac{1}{T}\mathcal{E}(\{\sigma_i\})} \tag{4.2}$$

A probability distribution as in (4.1) is called a *Gibbs measure*. It is customary to introduce the inverse temperature $\beta = \frac{1}{T}$ and to write (4.1) as:

$$p(\{\sigma_i\}) = \mathcal{Z}^{-1} e^{-\beta\,\mathcal{E}(\{\sigma_i\})} \tag{4.3}$$

There is a reason to introduce spin systems here: Obviously, any spin system can be interpreted as a voting system, we just interchange the words spin configuration and voting result as well as the symbols σ_i and X_i. In fact, a Gibbs measure p (as in (4.3)) defines a voting system with voting measure p, as long as $\mathcal{E}(\{\sigma_i\}) = \mathcal{E}(\{-\sigma_i\})$, so that p satisfies the symmetry condition (2.2). Moreover, any voting measure can be obtained from a Gibbs measure.

In particular, independent voting corresponds to the energy functional $\mathcal{E}(\{\sigma_i\}) \equiv 1$. In this case any configuration has the same energy, so that no configuration is more likely than any other.

The 'collective bias' model is given by an energy function:

$$\mathcal{E}(\{\sigma_i\}) = -h\sum_i \sigma_i \tag{4.4}$$

where h is a random variable connected to the collective bias variable Z by:

$$\frac{1}{2}(1+Z) = \frac{e^h}{e^h + e^{-h}}. \tag{4.5}$$

Note, that when h runs from $-\infty$ to ∞ in (4.5) the value of Z runs monotonously from -1 to $+1$.

In term of statistical physics in this model the spins do not interact with each other, but they do interact with a random but constant exterior field. The inverse temperature β is superfluous in this model as it can be absorbed in the magnetic field strength h.

5 The Voters' Interaction Model

In the collective bias model the voting behavior of each voter is influenced by a preassigned, a priori given collective bias variable Z (by an exterior magnetic field in the spin picture). The correlation between the voters results from the general voting tendency described by the value of Z.

In this section we investigate a model with a *direct* interaction between the voters, namely a tendency of the voters to vote in agreement with each other. In the view of statistical physics this corresponds to the tendency of magnets to align. There are various models in statistical physics to prescribe such a situation. Presumably the best known one is the Ising model where neighboring spins interact in the prescribed ways. The neighborhood structure is most of the time given by a lattice (e.g. $\mathbb{Z}^d$). The results on the system depend strongly on that neighborhood structure, in the case of the lattice $\mathbb{Z}^d$ on the dimension d.

In the following we consider another, in fact easier model where no such assumption on the local 'neighborhood' structure has to be made. We consider it an advantage of the model that very little of the microscopic correlation structure of a specific voting system enters into the model.

The model we are going to consider is known in statistical mechanics as the *Curie-Weiss model* or the *mean field model* (see e.g. Thompson 1972; Bolthausen and Sznitman 2002; Dorlas 1999). In this model a given voter (spin) interacts with all the other voters (resp. spins) in a way which makes it more likely for the voters (spins) to agree than to disagree. This is expressed through an energy function $\mathcal{E}$ which is smaller if voters agree. Note that a *small* energy for a given voting configuration (relative to the other configurations) leads to a *high* probability of that configuration relative to the others through formula (4.3).

The energy $\mathcal{E}$ for a given voting outcome $\{X_i\}_{i=1\ldots N}$ is given in the mean field model by:

$$\mathcal{E}\big(\{X_i\}\big) = -\frac{J}{N-1}\sum_{\substack{i,j\\ i\neq j}} X_i X_j. \tag{5.1}$$

Here J is a non negative number called the coupling constant. According to (5.1) the energy contribution of a single voter X_i is expressed through the averaged voting result of all other voters $\frac{1}{N-1}\sum_{j\neq i} X_j$. If X_i agrees in sign with this average the voter i makes a negative contribution to the total energy, otherwise X_i will increase the total energy. The strength of this negative or positive contribution is governed by the coupling constant J. In other words: Situations for which voter i

agrees with the other voters in average are more likely than others. This can be seen from the formula for the probability of a given voting outcome, namely:

$$p\left(\{X_i\}\right) = \mathcal{Z}^{-1}\, e^{-\beta\,\mathcal{E}(\{X_i\})} = \mathcal{Z}^{-1}\, e^{\beta J \frac{1}{N-1}\sum_{i\neq j} X_i X_j} \tag{5.2}$$

where as before

$$\mathcal{Z} = \sum_{\{X_i\}\in\{\pm 1\}^N} e^{-\beta\,\mathcal{E}(\{X_i\})}. \tag{5.3}$$

Since the probability density p depends only on the product of β and J we may absorb the parameter J into the inverse temperature β. So without loss of generality we can set $J = 1$. We denote the probability density (5.2) by $p_{\beta,N}$ and the corresponding expectation by $\mathbb{E}_{\beta,N}$.

Our goal is to compute the average:

$$w_N = \mathbb{E}_{\beta,N}\left(\,|\sum_{i=1}^{N} X_i\,|\,\right). \tag{5.4}$$

The quantity w_N gives the optimal weight in the council for a population of N voters with a correlation structure given by a mean-field model with inverse temperature β. We will see that the value of w_N changes dramatically when β changes from a value below one to a value above one. This has to do with the fact that the mean-field model undergoes a phase transition at the inverse temperature $\beta = 1$ (see Bolthausen and Sznitman 2002; Dorlas 1999; Thompson 1972).

Theorem 5.1.

(1) *If* $\beta < 1$ *then*

$$w_N = \mathbb{E}_{\beta,N}\left(\,|\sum_{i=1}^{N} X_i\,|\,\right) \sim \frac{\sqrt{2}}{\sqrt{\pi}}\,\frac{1}{\sqrt{1-\beta}}\,\sqrt{N} \quad \text{as } N \to \infty. \tag{5.5}$$

(2) *If* $\beta > 1$ *then*

$$w_N = \mathbb{E}_{\beta,N}\left(\,|\sum_{i=1}^{N} X_i\,|\,\right) \sim C(\beta) N \quad \text{as } N \to \infty. \tag{5.6}$$

Remarks 5.2.

(1) By $x_N \sim y_N$ as $N \to \infty$ we mean that $\lim_{n\to\infty} \frac{x_N}{y_N} = 1$.

(2) The constant $C(\beta)$ in (5.6) can be computed: If $\beta > 1$ then $C(\beta)$ is the (unique) positive solution C of

$$\tanh(\beta C) = C. \tag{5.7}$$

Note that for $\beta \le 1$ there is no positive solution of Eq. (5.7).

Theorem 5.1 can be understood quite easily on an intuitive level. We recall that the temperature T measures the strength of fluctuations, in other words: Low temperature (=large$\beta = \frac{1}{T}$) means high order in the system, high temperature (=smallβ) means disorder. Hence, the theorem says, that for strong order the expected voting result is well above (resp. well below) 50 % and the ideal weight is proportional to the size of the population, while for highly fluctuating societies polls are as a rule very tight and one obtains a square root law for the ideal representation.

The proof of Theorem 5.1 will be given in Sect. 10.

6 Conclusions

The above calculations show that one can reproduce the square-root law as well as the results of Gelman et al. (2004) and other laws by assuming particular correlation structures among the voters of a certain country. To find the right model is a question of adjusting the parameters of the models to empirical data of the country under consideration. Moreover, the models allow us to investigate questions about voting systems on a theoretical level. We believe that the models described above can help to understand voting behavior in many situations.

To *design* a nonhomogeneous voting system for a *constitution* in the light of our results is a question of different nature. Even knowing the correlation structure of the countries in question exactly would be of limited value to design a constitution. Constitutions are meant for a long term period, correlation structures of countries on the other hand are changing even on the scale of a few years.

One might argue that modern societies have a tendency to decrease the correlation between their members. In all modern states, at least in the West, the influence of churches, parties, and unions is constantly declining.

In addition to this it seems more important to protect small countries against a domination of the big ones than the other way round. This motivates us to choose a square-root law in these long term cases.

7 Appendix 1: Power Indices and Straffin's Model

Here we investigate some connection of our models with power indices. Power indices are usually defined through the ability of voters to change the voting result by their vote. To define power indices so we have to introduce a general setup for voting systems. This extends the considerations of the rest of this chapter where we considered only weighted voting. Our presentation below is inspired by Laruelle and Valenciano (2005) and Straffin (1982).

Let $\mathcal{V} = \{1, \ldots, N\}$ be the set of voters. The (microscopic) *voting outcome* is a vector $X = (X_1, \ldots, X_N) \in \{-1, +1\}^N$. Of course, $X_i = 1$ means that the voter i approves the proposal under consideration, while $X_i = -1$ means i rejects the proposal. We call $\Omega = \{-1, +1\}^N$ together with a probability measure $\mathbb{P}$ a *voting space*, if $\mathbb{P}$ is invariant under the transformation $T : \omega \mapsto -\omega$, thus $\mathbb{P}(\{X\}) = \mathbb{P}(\{-X\})$ (see Sect. 2 for a discussion of this property).

A *voting rule* is a function $\phi : \{-1, +1\}^N \longrightarrow \{-1, +1\}$. The voting rule associates to a microscopic voting outcome $X = (X_1, \ldots, X_N)$ a macroscopic voting result, i.e. the decision of the assembly. Thus, $\phi(X) = 1$ (resp. $\phi(X) = -1$) means that the proposal is approved (resp. rejected) by the assembly $\mathcal{V}$ if the microscopic voting outcome is $X = (X_1, \ldots, X_N)$. We always assume that the voting rule ϕ is *monotone*: If $X_i \leq Y_i$ for all i then $\phi(X) \leq \phi(Y)$. We also suppose that $\phi(-1, \ldots, -1) = -1$ and $\phi(+1, \ldots, +1) = +1$.

Following Laruelle and Valenciano (2005) we say that a voter i is *successful* for a voting outcome X if $\phi(X) = X_i$. Let us set $(X_1, \ldots, X_N)^{i,-} = (X_1, \ldots, X_{i-1}, -X_i, X_{i+1}, \ldots, X_N)$. We call a voter i *decisive* for X if $\phi(X) \neq \phi(X^{i,-})$, i.e. if the voting result changes if i changes his/her mind.

Given a voting space $(\{-1, 1\}^N, \mathbb{P})$ and a voting rule ϕ we define the ($\mathbb{P}$-) power index β by:

$$\beta(i) = \beta_{\mathbb{P}}(i) = \mathbb{P}\{X \in \{-1, +1\}^N \mid i \text{ is decisive for } X\} \tag{43}$$

Laruelle and Valenciano (2005) show that many known power indices are examples of the general concept (7.1). For example it is not difficult to show that one obtains the total Banzhaf index (see Banzhaf 1965 or Taylor 1995) if $\mathbb{P}$ is the probability measure of independent voting.

If we take $\mathbb{P} = \mathbb{P}_\mu$ to be the voting measure corresponding to the collective bias measure μ we get a whole family of power indices from (7.1). Straffin (1982) (see also Paterson 2012) demonstrates that if μ is the uniform distribution on $[-1, 1]$ then β_{P_μ} is just the Shapley-Shubik index (see Shapley and Shubik, M. 1954 or Taylor 1995).

In a subsequent publication on general power indices we will give a derivation of this fact in the current framework.

8 Appendix 2: Proofs for Section 2

We start with a short Lemma:

Lemma 8.1. *Suppose $X_1, \ldots, X_N$ are $\{-1, 1\}$—valued random variables with the symmetry property* (2.2) *then*

$$\mathbb{E}(\sum_{i=1}^{N} X_i) = 0 \tag{8.1}$$

and

$$\mathbb{E}\Big(\sum_{i=1}^{N} X_i \; \chi(\sum_{i=1}^{N} X_i)\Big) = \mathbb{E}\big(|\sum_{i=1}^{N} X_i|\big) \; . \tag{8.2}$$

Remarks 8.2 As defined above $\chi(x) = 1$ if $x > 0$, $\chi(x) = -1$ if $x \leq 0$.

Proof (2.2) implies

$$\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}$$

hence $\mathbb{E}(X_i) = 0$ and (8.1) follows.

To prove (8.2) we observe that

$$\begin{aligned}\mathbb{E}\Big(|\sum_{1}^{N} X_i|\Big) &= \mathbb{E}\Big(\sum_{i=1}^{N} X_i; \sum_{i=1}^{N} X_i > 0\Big) - \mathbb{E}\Big(\sum_{i=1}^{N} X_i; \sum_{i=1}^{N} X_i < 0\Big)\\ &= \mathbb{E}\Big(\sum_{i=1}^{N} X_i \; \chi(\sum_{i=1}^{N} X_i)\Big) \; .\end{aligned}$$

□

We turn to the proof of Theorem 2.1.

Proof (Theorem 2.1) Let us abbreviate: $S_\nu := \sum_{i=1}^{N_\nu} X_{\nu i}$.

Observe that the S_ν are independent by assumption and satisfy $\mathbb{E}(S_\nu) = 0$, moreover

$$\mathbb{E}\Big(S_\nu \; \chi(S_\mu)\Big) = 0 \text{ if } \nu \neq \mu \tag{8.3}$$

and

$$\mathbb{E}\Big(S_\nu \; \chi(S_\nu)\Big) = \mathbb{E}\big(|S_\nu|\big) \tag{8.4}$$

by Lemma 8.1. To find the minimum of the function

$$\Delta(w_1, \ldots, w_M) = \mathbb{E}((\sum_{1}^{M} S_\nu - \sum_{1}^{M} w_\nu \; \chi(S_\nu))^2)$$

we look at the zeros of $\frac{\partial \Delta}{\partial w_\mu}$.

$$0 = \frac{\partial \Delta}{\partial w_\mu} = -2\,\mathbb{E}\big((\sum_1^M S_\nu - \sum_1^M w_\nu\ \chi(S_\nu))\chi(S_\mu)\big)$$
$$= -2\,\mathbb{E}(S_\mu\ \chi(S_\mu) - w_\mu\ \chi(S_\mu)\chi(S_\mu))\ .$$

So

$$w_\mu\,\mathbb{E}((\chi(S_\mu))^2) = \mathbb{E}(S_\mu\ \chi(S_\mu)) = \ \mathbb{E}(|S_\mu|).$$

Since $\chi(S_\mu)^2 = 1$ we obtain

$$w_\mu = \ \mathbb{E}(|S_\mu|).$$

□

We turn to the proof of Theorem 2.2.

Proof Let $X_1,\ldots,X_N$ be $\{-1,1\}$—valued random variables with $P(X_i = 1) = P(X_i = -1) = \frac{1}{2}$. Then

$$\mathbb{E}(|\sum_1^N X_i|) = \sqrt{N}\,\mathbb{E}(|\frac{1}{\sqrt{N}}\sum_1^N X_i|).$$

By the central limit theorem (see e.g. Lamperti 1996) $\frac{1}{\sqrt{N}}\sum_1^N X_i$ has asymptotically a normal distribution with mean zero and variance 1, hence $\mathbb{E}(|\frac{1}{\sqrt{N}}\sum_1^N X_i|) \to \frac{\sqrt{2}}{\sqrt{\pi}}$. □

9 Appendix 3: Proofs for Section 3

Proof (Theorem 3.1) Since $\mathbb{E}_\mu(X_i) = 0$,

$$\begin{aligned}
\langle X_i, X_j\rangle_\mu &= \mathbb{E}_\mu(X_i X_j)\\
&= \mathbb{P}_\mu(X_i = X_j = 1) + \mathbb{P}_\mu(X_i = X_j = -1) - 2\mathbb{P}_\mu(X_i = 1, X_j = -1)\\
&= \int d\mu(\zeta)\{P_{\frac{1}{2}(1+\zeta)}(X_i = X_j = 1) + P_{\frac{1}{2}(1+\zeta)}(X_i = X_j = -1)\\
&\quad - 2P_{\frac{1}{2}(1+\zeta)}(X_i = 1, X_j = -1)\}\\
&= \int d\mu(\zeta)\{\frac{1}{4}(1+\zeta)^2 + \frac{1}{4}(1-\zeta)^2 - \frac{1}{2}(1-\zeta^2)\}\\
&= \int \zeta^2 d\mu(\zeta).
\end{aligned} \tag{9.1}$$

□

To prove Theorem 3.2 we need the following Lemma:

Lemma 9.1 $\mathbb{E}_\mu(\frac{1}{N}|\sum(X_i - Z)|) \leq \frac{1}{\sqrt{N}}$.

Proof

$$\begin{aligned} \mathbb{E}_\mu\big(\frac{1}{N}|\sum(X_i - Z)|\big) &= \frac{1}{N}\,\mathbb{E}_\mu\big(|\sum(X_i - Z)|\big) \\ &\leq \frac{1}{N}\Big\{\,\mathbb{E}_\mu\Big(\big(\sum(X_i - Z)\big)^2\Big)\Big\}^{1/2} \\ &= \frac{1}{N}\,\Big\{\int d\mu(\zeta)\; E_{P_\zeta}\Big(\big(\sum_1^N(X_i - \zeta)\big)^2\Big)\Big\}^{1/2}. \end{aligned} \tag{9.2}$$

Given $Z = \zeta$ the random variables $X_i - \zeta$ have mean zero and are independent with respect to the measure P_{p_ζ}, thus

$$E_{P_\zeta}\Big(\big(\sum_1^N(X_i - \zeta)\big)^2\Big) = NE_{P_\zeta}(X_i - \zeta)^2 = N(1-\zeta^2) \leq N,$$

hence

$$(9.2) \leq \frac{1}{\sqrt{N}}\big(\int d\mu(\zeta)(1-\zeta^2)\big)^{1/2} \leq \frac{1}{\sqrt{N}}.$$

□

Using Lemma 9.1 we are in a position to prove Theorem 3.2:

Proof (1) Suppose that:

$$\overline{\mu}_N = \int |\zeta|\; d\mu_N(\zeta) \geq C\,\frac{1}{N^{1/2-\varepsilon}} \tag{9.3}$$

then we estimate:

$$\begin{aligned} \mathbb{E}_{\mu_N}\big(\frac{1}{N}|\sum_1^N X_i|\big) &= \mathbb{E}_{\mu_N}\big(|\frac{1}{N}\sum_1^N(X_i - Z) + Z|\big) \\ &\leq \mathbb{E}_{\mu_N}(|Z|) + \mathbb{E}_{\mu_N}\big(|\frac{1}{N}\sum_1^N(X_i - Z)|\big) \\ &\leq \overline{\mu}_N + \frac{1}{\sqrt{N}} \end{aligned} \tag{9.4}$$

by Lemma 9.1. Moreover

$$\begin{aligned}\mathbb{E}_{\mu_N}(\frac{1}{N}|\sum_1^N X_i|) &\geq \mathbb{E}_{\mu_N}(|Z|) - \mathbb{E}_{\mu_N}(|\frac{1}{N}\sum X_i - Z|)\\ &\geq \overline{\mu}_N - \frac{1}{\sqrt{N}}.\end{aligned} \tag{9.5}$$

Hence

$$|\,\mathbb{E}_{\mu_N}(\frac{1}{N}|\sum_1^N X_i|) \;-\; \overline{\mu}_N\,| * \leq * \frac{\varnothing}{\sqrt{}} \tag{9.6}$$

which proves (3.12).

(2) To prove (3.14) we obtain by the same reasoning as above:

$$|\mathbb{E}_{\mu_N}(\frac{1}{N}|\sum\nolimits_1^N X_i|) - \overline{\mu}_N| \leq \frac{1}{\sqrt{N}} \tag{9.7}$$

□

We end this section with the proof of Theorem 3.6:

Proof We have to prove that for bounded continuous functions f:

$$\int \Big(f(\frac{1}{N}\sum_{i=1}^N X_i) - f(Z)\Big)\, d\,\mathbb{P}_{\mu_N} * \to * \gtrless. \tag{9.8}$$

The convergence (9.8) is clear for continuously differentiable f from Lemma 9.1. It follows for arbitrary bounded continuous f by a density argument. □

10 Appendix 4: Proofs for Section 5

In this section we prove Theorem 5.1.

Proof (Theorem 5.1 (1))

We denote by $E_0^{(N)}$ the expectation of the coin tossing model for N independent symmetric $\{+1, -1\}$-valued random variables, i.e.:

$$E_0^{(N)}\big(F(X_1,\ldots,X_N)\big) = \frac{1}{2^N}\sum_{x_i\in +1,-1^N} f(x_1,\ldots,x_N). \tag{10.1}$$

We set:

$$\mathcal{Z}_{\beta N} = E_0^{(N)}\Big(\, e^{\frac{\beta}{2}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N X_i\right)^2}\Big) \tag{10.2}$$

and:

$$\mathcal{X}_{\beta N} = E_0^{(N)}\Big(\,|\frac{1}{\sqrt{N}}\sum_{i=1}^{N} X_i|\; e^{\frac{\beta}{2}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N} X_i\right)^2}\Big). \tag{10.3}$$

Then:

$$\mathbb{E}_{\beta N}(|\sum_{i=1}^{N} X_i|) = \sqrt{N}\,\frac{\mathcal{X}_{\beta,N}}{\mathcal{Z}_{\beta,N}}. \tag{10.4}$$

Under the probability law $E_0^{(N)}$ the random variables X_i are centered and independent, thus the central limit theorem (see e.g. Lamperti 1996) tells us that $\frac{1}{\sqrt{N}}\sum_{i=1}^{N} X_i$ converges in distribution to a standard normal distribution. Consequently, for $\beta < 1$ and $N \to \infty$:

$$\mathcal{Z}_{\beta N} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(1-\beta)x^2}{2}}\, dx = \frac{1}{\sqrt{1-\beta}} \tag{10.5}$$

and:

$$\mathcal{X}_{\beta N} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |x| e^{-\frac{(1-\beta)x^2}{2}} dx = \frac{\sqrt{2}}{\sqrt{\pi}}\frac{1}{1-\beta}. \tag{10.6}$$

Consequently:

$$\mathbb{E}_{\beta N}(|\sum_{i=1}^{N} X_i|) = \sqrt{N}\,\frac{\mathcal{X}_{\beta,N}}{\mathcal{Z}_{\beta,N}} \sim \frac{\sqrt{2}}{\sqrt{\pi}}\frac{1}{\sqrt{1-\beta}}\sqrt{N}. \tag{10.7}$$

□

Proof (Theorem 5.1 (2))

By Theorem 6.3 in Bolthausen and Sznitman (2002) the distribution ν_N of $S_N = \frac{1}{N}\sum_{i=1}^{N} X_i$ converges weakly to the measure $\nu = \delta_{-C(\beta)} + \delta_{C(\beta)}$ where $C(\beta)$ was defined in (5.7).

Hence,

$$\mathbb{E}_{\beta}(|\sum_{i=1}^{N} X_i|) = N\mathbb{E}_{\beta}(|S_N|) \tag{10.8}$$

$$= N \int |\lambda| d\nu_N(\lambda) \tag{10.9}$$

$$\approx N \int |\lambda| d\nu(\lambda) \tag{10.10}$$

$$= NC(\beta). \tag{10.11}$$

Acknowledgments It is a pleasure to thank Hans-Jürgen Sommers, Duisburg-Essen, and Wojciech Słomczyński and Karol Życzkowski, Krakow, for valuable discussions. The author would also like to thank Helmut Hölzler, Göttingen, for pointing out a notational inconsistency in a previous version of the chapter.

References

Banzhaf, J. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review*, *19*, 317–343.

Bolthausen, E., & Sznitman, A. (2002). Ten lectures on random media. Basel: Birkhäuser.

Dorlas, T. (1999). *Statistical mechanics*. Bristol: Institute of Physics Publishing.

Felsenthal, D., & Machover, M. (1998). *The measurement of power: Theory and practice, problems and paradoxes*. Cheltenham: Edward Elgar.

Gelman, A., Katz, J., & Bafumi, J. (2004). Standard voting power indexes don't work: An empirical analysis. *British Journal of Political Science*, *34*, 657–674.

Kirsch, W. (2012) *On the distribution of power in the Council of ministers of the EU*. http://www.ruhr-uni-bochum.de/mathphys/.

Lamperti, J. (1996). Probability. New York: Wiley.

Laruelle, A., & Valenciano, F. (2005). Assessing success and decisiveness in voting situations. *Social Choice and Welfare*, *24*, 171–197.

Paterson, I. (2012). *A lesser known probabilistic approach to the Schapley-Shubik index and useful related voting measures* (Preprint).

Shapley, L., & Shubik, M. (1954). A method for evaluating the distribution of power in a commitee system. *American Political Science Review*, *48*, 787–792.

Straffin, P. (1982). Power indices in politics. In S.J. Brams et al. (Eds.), *Political and related models*. Heidelberg: Springer.

Taylor, A. (1995). *Mathematics and politics*. New York: Springer.

Thompson, C. (1972). *Mathematical statistical mechanics*. Princeton: Princeton University Press.

A New Analysis of a Priori Voting Power in the IMF: Recent Quota Reforms Give Little Cause for Celebration

Dennis Leech and Robert Leech

1 Introduction

The governance of the Bretton Woods institutions (the International Monetary Fund and the World Bank) is by means of a system of weighted voting. All member countries have a voice but cast different numbers of votes depending on their quotas (the IMF term for the country's financial stake), or their shareholdings (the term used by the World Bank). In the IMF each country's number of votes is determined by a formula that gives it a number of so-called basic votes that each country has plus a number proportional to its quota.[1] The rules require that all countries that are members of the World Bank must also be members of the IMF, and their shareholdings depend very strongly on their IMF quotas. It is therefore not necessary to make separate analyses of voting power for the two institutions and our findings about the IMF can therefore be taken as broadly applying to the World Bank also.

Weighted voting in the IMF is problematic because it results in a severe democratic imbalance with the distribution of voting power being massively biased against the developing and poor countries (Buira 2002). This dominance by the industrialized countries has been criticised by the developing countries and others as leading the organization to adopt policies that have taken insufficient

This study updates and modifies "Voting Power in Bretton Woods Institutions" in Power Measures III [Homo Oeconomicus 22(4)], edited by Gianfranco Gambarelli and Manfred J. Holler, Munich: Accedo Verlag, 2005: 605–627.

[1] One vote for every 100,000 special drawing rights of quota. Each country's quota is its financial stake in the IMF and theoretically meant to reflect its importance in the world economy.

D. Leech (✉)
Department of Economics, University of Warwick, Coventry CV4 7AL, UK
e-mail: d.leech@warwick.ac.uk
URL: http://www2.warwick.ac.uk/fac/soc/economics/staff/academic/leech

R. Leech
Imperial College, London, UK

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*, DOI: 10.1007/978-3-642-35929-3_21, © Springer-Verlag Berlin Heidelberg 2013

notice of the interests of those countries, especially the imposition of conditions on borrowers derived from extreme neo-liberal economics in the so-called Washington Consensus. (See e.g. Buira 2003; Woods 2006). The need for reform of the governance of the international financial organisations was accepted by all countries as part of the Monterrey Consensus agreed in 2002 (Buira 2006) and again at the follow-up meeting in Doha in 2008. Changes aimed at giving greater voice to poor countries and emerging economies have now started to be implemented and it is of interest to study how effective they are.

Besides this inequality in the voting arrangements—inequality that has historically been intended as part of the design of an institution on the principle that those who contribute most should have the most say—there exists a *further* bias, resulting from the use of weighted voting. The idea of weighted voting is that each country's voting power should be predetermined and that it should be proportional to its voting weight. However, a member's voting *power* is not the same as its *weight*: its *power* is its ability to be decisive whenever a vote is taken—to make a difference to the outcome—whereas its *weight* is just the number of votes it has been allocated by the rules.

It follows that voting power is a fundamental property of the rules by which decisions are taken, together with the weights of all voters, and this can only be revealed by detailed analysis that looks at outcomes, using voting power indices, Because this important distinction is often ignored in practice, designing constitutions that use weighted voting often leads to undesired or unexpected consequences in terms of the distribution of voting power among countries.

The voting weights in the IMF are very unequal: the USA has more than two and a half times as many votes as the country with the next-largest voting weight, Japan. We use power indices to measure each member's voting power. The USA turns out to have much more voting power than weight. This disproportionality is another argument for reforming the weights in a more radical direction than has hitherto been suggested. More generally the lack of a direct link between power and weight adds to the case for decoupling the allocation of votes from both the provision of and access to finance.

Defenders of the present voting system claim it embodies democratic accountability if one accepts the principle that voting rights should be attached to the supply of capital. For example, when he was Managing Director Horst Köhler said: "I would also like to underline that still we are a financial institution, and a financial institution means you need also to have someone who provides capital and I think there is a healthy element in the fact that the provision of capital and voting rights is, in a way, combined, because this is also an element of efficiency, of accountability."[2] The distorting effect of weighted voting that we describe below makes this claim far from being true.

As a general principle weighted voting is an attractive idea because it offers the prospect of designing an intergovernmental decision-making body that could have

---

[2] House of Commons Treasury Select Committee, 4th July 2002.

a real claim to democratic legitimacy—for example, in an institution of world government where a country's voting power reflects its population. But it is important to be clear about what we mean by weighted voting. Systems based on the use of a *bloc vote* where a country (or group of countries acting together) casts all its voting weight as a single unit, as in the IMF, cannot be relied on to work like that and in general they do not, as we will show. On the other hand, if the rules are such that a country is represented by a *number of delegates* each with one vote that they are allowed to cast individually, rather than having to vote together as a unit, then this problem does not arise. The latter is simply a representative democracy and the number of votes or delegates is equivalent to the country's power. The argument we are advancing here is only relevant when the votes cannot be split.

We will use the method of voting power analysis to explore the relationships between the voting weights, the decision rule and the resulting voting powers of the members. This requires us to analyse all the voting outcomes that can occur, and investigate the ability of every member to be decisive, to be able to decide whether the vote leads to a decision or not. We will use voting power indices to compare the powers of different members.

Our principal finding is that the voting power of the USA is far greater than its voting weight. That is, its actual power over decision-making far exceeds its nominal voting power. We also use the method for two important analyses: first the effect of the *ad hoc* increase in voting weight that occurred in 2008 for four emerging economies (China, Korea, Mexico and Turkey) that were previously very badly unrepresented, second the more radical reforms agreed at the Singapore meeting in 2008. Secondly we consider the Executive Board as a representative body in which the directors are elected by constituencies of countries by majority voting. We find that the constituency system considerably enhances the power of certain smaller European countries, especially Belgium, the Netherlands and Switzerland.

We begin with an outline of the principles of voting power analysis in the next section. Then in Sect. 3 the system of governance of the IMF is described, in Sects. 4 and 5 we present the analyses of the Board of Governors and the Executive Directors, in 6 we consider the voting power implications for treating the constituency system that underpins the Executive Board as form of democratic representation assuming formal voting within constituencies.

2 Weighted Voting and Voting Power Analysis

A country's voting power is its potential to be decisive in a decision taken by vote, measured by the probability with which it can change what would otherwise be a losing vote to a winning one. In general this has a rather imprecise relation with its weight. In reality its power depends on all the other members' weights as well as the voting rule by which decisions are taken. A case that shows the issue starkly is that of the European Economic Community which also employed weighted voting in the council of ministers: the distribution of voting power among the six members was far from proportional to voting weight between 1958 and 1972. See Leech and Leech (2005b) for the details. (Brams and Affuso 1976, were the first to show it).

By considering all possible voting outcomes the method of power indices is technically that of a priori voting power: each member's power index is its decisiveness as a fraction of the theoretically possible outcomes without regard for the likelihood of their occurring. The method can be thought of as an analysis of the implications for power of the voting rules, considered in the abstract, as giving what can be called constitutional power.[3] Probability calculus is used as a tool for calculating the power indices. Technically the probability of a voter being decisive is the Penrose index (also known equivalently as the Penrose measure, Penrose-Banzhaf index, Absolute Banzhaf index). This is a measure of the a priori probability of the voter being decisive and is the simplest index for the purpose. Other power indices could be used, but we take the view that the superiority of the Penrose-Banzhaf index is established on both theoretical and empirical grounds. See Felsenthal and Machover (1998); Leech (2002c) for a comparison with the other so-called classical power index, the Shapley-Shubik; see also Coleman (1971). However, since our purpose is to investigate changes in relative voting power among the member countries, we use the normalized version, generally known as the Banzhaf index (or Normalised Banzhaf index), that has the property that the indices over all the voters sum to one, and therefore it provides a distribution of voting power. We will refer to values of this index as voting powers.

Voting power analysis will be used in two ways. First it will be used to analyse power relations in the existing governance structures of the IMF, the Board of Governors and the Executive Board, and also the effects of recent reforms. These will be the main empirical results of the paper.

Second, we also use it to study the properties of indirect or two-level voting procedures implied by the IMF constituency system where countries are placed in a series of groups, each containing a number of members, where each group's Executive Director casts all its members' votes *en bloc* in the second stage the Executive. The Penrose index described above provides a simple methodology for doing such analysis, since any member's (indirect) power index is simply obtained as the product of the two relevant power indices in the two stages, each of which is an independent probability. These absolute indices are then normalized to sum to one as before to provide a distribution of indirect voting power for this voting body. The theory is described in Leech and Leech (2005a, 2006); the method in terms of game theory is presented in Owen (1995).

This method follows that proposed in Coleman (1973) to address the general question of why social actors give up power to join groups. By joining with others

[3] No consideration is given here for the members' preferences, which would determine the likelihood of particular members voting in the same way as each other, which would produce an analysis of empirical voting power. Some coalitions look less likely than others from a game-theoretic point of view. Such an analysis is beyond the scope of the present study (see Leech 2003b).

in a group, an actor gives up his power as an independent voter in favour of more limited power over group decisions, but may nevertheless gain overall if the group's 'power of combined forces' is sufficiently greater that it offsets that loss of power. The use of power indices permits results to be obtained very easily since it allows us to combine the power of the actor within the group and the power of the group. Analytically this can be thought of as equivalent to a compound voting game. This approach lends itself naturally to the analysis of intergovernmental weighted voting where there is accountability to a lower body, such as a country's electorate, parliament or a regional intergovernmental grouping. It can be generalized to compound voting games with three or more levels. It is a useful tool for the analysis of voting power implications of changes to the architecture of voting in the international institutions. We emphasise that such scenarios are very stylized and open to criticism for their realism.

3 Weighted Voting in the IMF

All countries are members of the Board of Governors, and as such have direct representation at the highest level of formal decision-making, but the real management is done by the Executive Directors (also known as the Executive Board).

In the Board of Governors and in the election of Executive Directors the voting weight of each country is made up of two components: a fixed component of so-called 'basic' votes which is the same for each country, and a variable component that depends on the country's quota. This formula for determining voting weight is intended as a compromise between two principles: the equal representation of member countries (via the basic votes), analogous to the UN General Assembly, and voting power based on contributions in the manner of a joint stock company. Over time the basic element has become severely eroded and the quota, or share-based votes, have become dominant. This is an important factor behind the disempowerment of the poor countries. The restoration of the basic votes to their original level is a main aim of the reform movement.

There are currently (in 2012) 188 members, of whom the USA has by far the largest voting weight, with 421,965 votes, 16.75 % of the total, and the smallest is Tuvalu with 759 votes, 0.03 %. The second-highest voting weight is held by Japan with 6.23 %, Germany 5.81 %, France and UK with 4.29 % and so on.

The Executive Board consists of 24 members, some of whom are appointed by their governments and some elected by member states. Five directors are appointed by the members with the largest quotas: USA, Japan, Germany, France and UK. The remaining 19 directors are elected.

In meetings when a vote is taken on an issue, the Executive Board uses weighted voting exactly like the Board of Governors: the appointed directors cast the number of votes of the member that appointed them, and the elected directors cast the combined number of votes of the countries that voted for them. There is no provision for executive directors to split their vote to reflect the views of the countries that voted for them, when they are not unanimous on the issue, although they are allowed to abstain, which in a sense is equivalent to splitting their voting weight equally and voting for both sides.

There are elections for directors every two years. Each eligible member country votes for a single director and directors are elected in order of the number of votes they receive. The rules for electing directors lay down a minimum and maximum number of votes that can be cast for each elected director, and hence sizes of the weighted votes that they can cast at Executive Board meetings which prevent any elected director becoming too powerful. Eliminating ballots are taken until all the vacant directorships are filled.[4]

The result is a pattern of voting power generally similar to that of the governors.[5] Three directors are elected by a single country, so are in effect appointed: China, Russia and Saudi Arabia. The rest are elected by groups of countries.

A variety of decision rules are used for different types of decisions. Ordinary decisions are made by simple (weighted) majority of the votes cast (the quorum for meetings of the Board of Governors being a majority of members having not less than two-thirds of the voting weight; that for the Executive Board being a majority of directors having not less than one-half of the total voting weight). A number of matters require decisions to be taken by a supermajority of 85 %. This supermajority, taken in conjunction with the weight of the USA, 16.75 %, means that the USA is the only member that possesses a veto.

It is customary for official spokespersons to say that decisions in the Executive are normally taken by consensus and formal votes are avoided. However, this claim is not universally accepted, many writers pointing out that the absence of formal voting is not the same thing as consensus decision making. In practice decision making during a debate where there is contention involves the secretary informally keeping a tally of the weighted votes held by the executive directors who speak on each side according to the sense of their contribution, a 'consensus' being deemed to have been found when the required majority has been reached. Thus, although a formal vote is usually avoided, the rule may be closer to weighted majority voting than consensus building. See Buira (2005); Woods (2001).

The American veto has always been an important aspect of the governance of the institutions, and continues to be so, the articles having been amended to increase the supermajority threshold for special decisions from 80 to 85 % when the USA decided to reduce its quota. The existence of this veto power does not mean that the USA can be said to *control* the institutions, however. On the contrary,

---

[4] The rules are laid down in Schedule E of the Articles of Association. They state that, in order to be elected, a director must receive at least four percent and no more than nine percent of the eligible votes. If the number of directors elected by this procedure is less than the number required, then there are further ballots with voting eligibility restricted to (1) those members who voted for a candidate who received less than 4 % and (2) those members who voted for a director who was elected but whose votes are deemed to have taken the votes for the director above 9 %.

[5] The voting weights as proportions vary slightly between the Board of Governors and Executive Directors because of differences in participation in votes by countries with small weights.

although it gives it absolute unilateral blocking power, at the same time it also limits that country's power because it equally ensures a collective veto for small groups of other countries. Formally, in terms of Coleman's terminology, while the supermajority rule gives the United States complete *power to prevent action*, it also limits its *power to initiate action* (Coleman 1971). Therefore, its voting power—and its power index (which is an average of these two)—is limited. The existence of the 85 % supermajority can be seen to give veto power to three other countries acting together (for example, Japan, Germany and France). The developing countries, if they acted as a bloc, or the EU countries, or many other similar small groups, obviously have a veto.[6] The 85 % rule tends to equalize voting power. Taking the argument to its limit, the case of a unanimity rule (i.e. a supermajority requirement of 100 %) would give every member a veto and equalise power, making voting weight irrelevant. For these reasons the power analysis in this study considers only ordinary decisions that require a simple majority vote. Analysis of power under supermajorities has been made in Leech (2002a).

4 Voting Power in the Board of Governors

Table 1 shows three analyses which reveal the weighted voting effect and give a picture of the effects of the quota reforms:

(1) for 2006 before the reforms;
(2) for 2008 after the *ad hoc* adjustments to the quotas for four emerging economies that were seriously out of line: China, Korea, Mexico and Turkey; and
(3) for 2012, after the partial implementation[7] of the reforms agreed in 2010.

The table shows, for each of the main countries, (1) its relative voting weight and (2) its normalized power index or vote share,[8] in each of the years. Significant changes in weights in the reforms are highlighted in bold. The table also shows the Gini coefficient of inequality for both the voting weights and the voting power

[6] This point about the difference between veto power and the power of control was made very clearly by Keynes in opposition to the proposed American veto based on supermajorities in a speech to the House of Lords in 1943 when the Bretton Woods institutions were being planned. See Moggeridge (1980), p. 278; also his Letter to J. Viner, p. 328. Keynes advocated simple majority voting.

[7] Quotas change when countries make the payments, which not all have done at the time of writing.

[8] These power indices have been calculated using the computer program *ipmmle* (accessible online at www.warwick.ac.uk/~ecaae, Leech and Leech 2003) which implements the algorithm for computing power indices for voting bodies that are large both in having many members and where the voting weights are large numbers, described in Leech (2003a). For an overview of computing power indices see Leech (2002b) See also Leech (2011) for the properties of power indices when the number of voters in very large.

Table 1 Voting weights and voting powers in the Board of Governors (largest weight countries)

2006			2008			2012		
	Weight	Power		Weight	Power		Weight	Power
USA	17.09	24.49	USA	16.77	23.80	USA	16.75	24.29
Japan	6.13	5.46	Japan	6.02	5.41	Japan	6.23	5.50
Germany	5.99	5.35	Germany	5.88	5.30	Germany	5.81	5.18
France	4.95	4.48	France	4.86	4.42	UK	4.29	3.89
UK	4.95	4.48	UK	4.86	4.42	France	4.29	3.89
Italy	3.25	2.97	China	3.66	3.35	China	3.81	3.46
Saudi	3.22	2.94	Italy	3.19	2.93	Italy	3.16	2.88
China	2.94	2.69	Saudi	3.16	2.90	Saudi	2.80	2.56
Canada	2.94	2.69	Canada	2.89	2.65	Canada	2.56	2.34
Russia	2.74	2.50	Russia	2.69	2.47	Russia	2.39	2.18
Netherland	2.38	2.18	Netherland	2.34	2.15	India	2.34	2.14
Belgium	2.13	1.95	Belgium	2.09	1.92	Netherland	2.08	1.90
India	1.92	1.76	India	1.89	1.74	Belgium	1.86	1.70
Switzerland	1.60	1.46	Switzerland	1.57	1.45	Brazil	1.72	1.57
Australia	1.50	1.37	Australia	1.47	1.35	Spain	1.63	1.49
Spain	1.41	1.29	Mexico	1.43	1.32	Mexico	1.47	1.35
Brazil	1.41	1.29	Spain	1.39	1.28	Switzerland	1.40	1.28
Venezuela	1.23	1.13	Brazil	1.38	1.27	Korea	1.37	1.25
Mexico	1.20	1.10	Korea	1.33	1.23	Australia	1.31	1.20
Sweden	1.11	1.02	Venezuela	1.21	1.11	Venezuela	1.09	1.00
Argentina	0.98	0.90	Sweden	1.09	1.01	Sweden	0.98	0.90
Indonesia	0.97	0.89	Argentina	0.97	0.89	Argentina	0.87	0.80
Austria	0.87	0.80	Indonesia	0.95	0.87	Austria	0.87	0.80
South Africa	0.87	0.80	Austria	0.86	0.79	Indonesia	0.86	0.79
Nigeria	0.82	0.75	South Africa	0.85	0.79	Denmark	0.78	0.71
Norway	0.78	0.71	Nigeria	0.80	0.74	Norway	0.78	0.71
Denmark	0.77	0.71	Norway	0.77	0.70	South Africa	0.77	0.71
Korea	0.76	0.70	Denmark	0.75	0.69	Malaysia	0.73	0.67
Iran	0.70	0.64	Iran	0.69	0.63	Nigeria	0.73	0.67
Malaysia	0.69	0.63	Malaysia	0.68	0.63	Poland	0.70	0.64
Kuwait	0.65	0.60	Kuwait	0.63	0.58	Iran	0.62	0.57
Ukraine	0.64	0.59	Ukraine	0.63	0.58	Turkey	0.61	0.56
Poland	0.64	0.59	Poland	0.63	0.58	Thailand	0.60	0.55
Finland	0.59	0.54	Finland	0.58	0.54	Singapore	0.59	0.54
Algeria	0.59	0.54	Algeria	0.58	0.53	Kuwait	0.58	0.53
Iraq	0.56	0.51	Turkey	0.55	0.51	Ukraine	0.57	0.52
Libya	0.53	0.49	Iraq	0.55	0.50	Finland	0.53	0.49
Thailand	0.51	0.47	Libya	0.52	0.48	Ireland	0.53	0.49
Hungary	0.49	0.45	Thailand	0.50	0.46	Algeria	0.53	0.49
Pakistan	0.49	0.45	Hungary	0.48	0.44	Iraq	0.50	0.46
Romania	0.49	0.45	Pakistan	0.48	0.44	Libya	0.48	0.44
Turkey	0.45	0.41	Romania	0.48	0.44	Greece	0.47	0.43
…	…	…	…	…	…	…	…	…
	100.00	100.00		100.00	100.00		100.00	100.00
Gini	0.7780	0.7958		0.7819	0.7990		0.7584	0.7767

Power indices have been calculated using the program *ipmmle* available on the website www.warwick.ac.uk/~ecaae

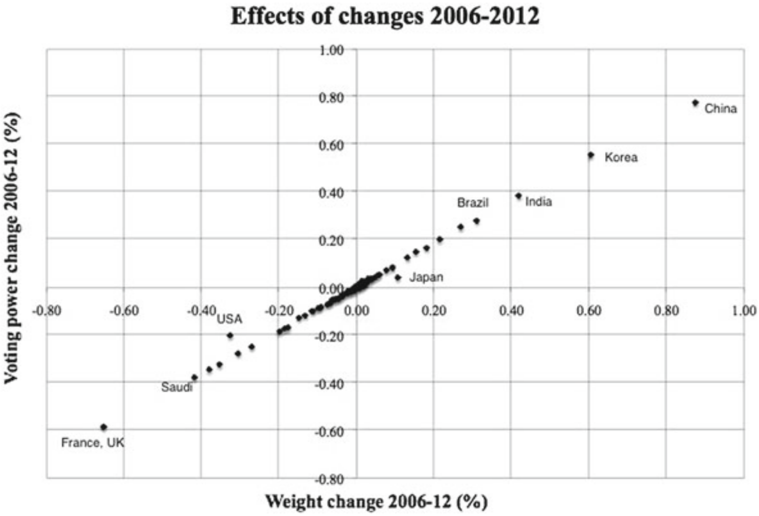


Fig. 1 Voting power implications of the quota reforms

indices. Inequality is very high in 2006 and the reforms reduce it by very little. Inequality in voting power is slightly higher than it is for weight.

The table shows that the voting power of the United States is considerably out of line with its weight. In 2006 its voting weight of just over 17 % gave it 24 % of the voting power. Its weight went down slightly in 2008 and again in 2012 but it was still massively dominant giving it much greater voting power. All other members have less power than their weight. Thus, we can say that the weighted voting system has a hidden tendency to enhance the power of the USA at the expense of all other countries.

The 2006 table also brings out a number of glaring anomalies pointing to the need for reform. Canada and China had the same number of votes, and voting power, despite the economy of China being much bigger than that of Canada. This bias against developing countries is seen, also, in the comparison of the voting weight of some rich countries like Belgium, Netherlands and Spain with large emerging economies especially India, Brazil and Mexico. A particularly glaring juxtaposition is that between Denmark and Korea, the former having more voting weight than the latter despite its economy being much smaller.

The implications of the quota reforms are also illustrated in Fig. 1 which shows the changes in voting power indices plotted against the changes in weights. The reforms have been in two stages: first the *ad hoc* increases for China, Korea, Mexico and Turkey implemented in 2008, then the changes resulting from the more radical reforms implemented in 2012. The latter reforms were: (1) the introduction of a more transparent, simpler formula to replace the previous complicated five-fold system; (2) tying the quotas more closely to the formula; (3) tripling of basic votes for all members; (4) a second round of *ad hoc* increases for the four countries mentioned above. This second round of reforms was accompanied by an increase in general quotas. The main changes in relative voting weights were increases for China, Korea, India, Brazil, Mexico and some others at the expense of the USA, some European countries notably the UK and France, Saudi Arabia and Canada. None of the changes was greater than one percent of the total voting weight, so perhaps it is not surprising that the voting power effects are very small.

Figure 1 shows a common pattern for all countries, except the two with the largest quotas. For all countries except the USA and Japan, the voting power change is proportional to the weight change; they all lie on a straight line through the origin with a gradient of less than 1. Those countries whose weight increases gain slightly less voting power while those whose weight falls lose less voting power. China gains 0.87 % in weight but only 0.78 % in voting power, while the UK and France lose 0.65 % weight but only 0.58 % voting power. The exceptions are the USA and Japan which lie above and below the line respectively. Japan gains 0.1 % in weight, but only 0.04 % in voting power. The United States loses 0.32 % weight but only 0.2 % in voting power; it loses about the same voting weight as the Netherlands (0.30 %) but does not lose comparable voting power (0.28 %). These are unexpected weighted voting effects due to the great inequality in weights.

However, these effects are all small. They provide little to support the claim of the then IMF Managing Director, Dominique Strauss-Kahn that, "Taken together, it's a big shift in quotas and accordingly in voting power. It's a very important increase in the voice and representation of the emerging market and developing countries. It is a historical reform of the IMF."[9]

Figure 2 shows the changes in voting power and weights that followed the *ad hoc* quota increases for the four countries implemented in 2006. These changes did not involve changes to the voting weights of the other members. The main effect was that the USA lost weight and power while China, Korea, Mexico and Turkey all gained. Interestingly the USA lost more in power (0.69 %) than in weight (0.31 %). The second phase of the reforms are shown in Fig. 3, which compares 2008 with 2012. Now we see that the USA gained in voting power (0.5 %) as a result of the reforms although its relative weight hardly changed at all. The biggest gainers from this phase were India, Brazil and Spain, while the biggest losers were the UK and France.

[9] Press Release: "IMF Board Approves Far-Reaching Governance Reforms", 5 November 2010, IMF Washington.

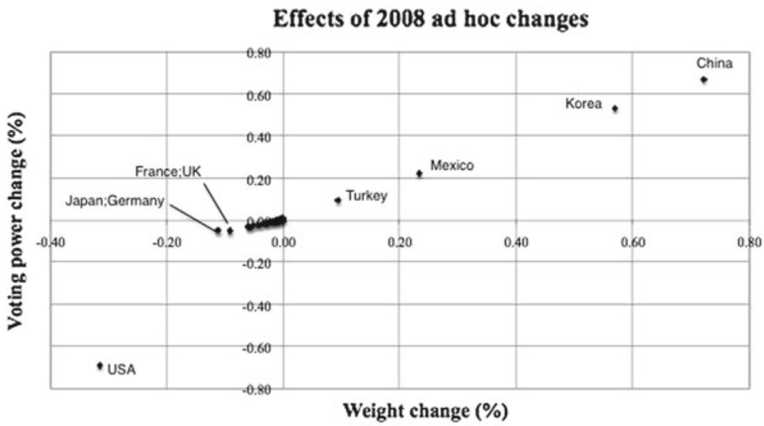


Fig. 2 The ad hoc increases in quota for China, Korea, Mexico and Turkey

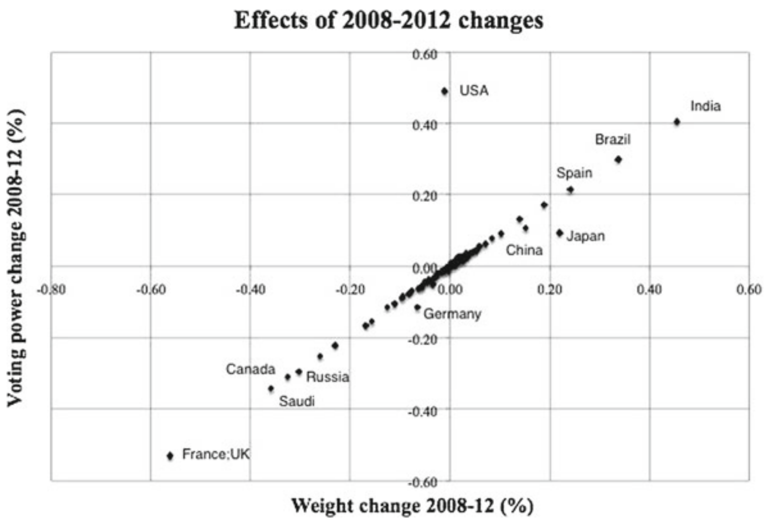


Fig. 3 The second phase of the quota reforms in 2012

5 Voting Power in the Executive Board

Table 2 shows the analyses, for 2006 and 2012,[10] for the Executive Board, which has 24 executive directors who cast weighted votes. The directors of the countries with the biggest five quotas (USA, Japan, Germany, UK and France) together with those of China, Russia and Saudi Arabia, are directly appointed by their

[10] We omit 2008 because the changes were so small.

Table 2 Voting weights and voting powers in the Executive Directors

	2006					2012				
		No.	Voting weight (%)	Voting power (%)	Ratio		No.	Voting weight (%)	Voting power (%)	Ratio
1	USA	1	17.09	21.45	1.255	USA	1	16.79	20.95	1.248
2	Japan	1	6.13	5.82	0.949	Japan	1	6.25	5.95	0.952
3	Germany	1	5.99	5.69	0.949	Germany	1	5.82	5.55	0.952
4	Belgium	10	5.13	4.87	0.948	Belgium	10	4.98	4.74	0.951
5	France	1	4.95	4.69	0.948	Mexico	8	4.66	4.43	0.951
6	UK	1	4.95	4.69	0.948	Netherland	13	4.53	4.30	0.950
7	Netherland	12	4.85	4.59	0.948	France	1	4.30	4.09	0.950
8	Mexico	8	4.27	4.05	0.948	UK	1	4.30	4.09	0.950
9	Italy	7	4.18	3.96	0.948	Italy	7	4.27	4.06	0.951
10	Canada	12	3.71	3.51	0.947	Singapore	13	3.94	3.75	0.950
11	Finland	8	3.51	3.32	0.947	China	1	3.82	3.63	0.949
12	Korea	14	3.33	3.15	0.947	Australia	15	3.63	3.45	0.950
13	Egypt	13	3.26	3.08	0.947	Canada	12	3.61	3.43	0.950
14	Saudi Arabia	1	3.22	3.05	0.946	Denmark	8	3.41	3.23	0.949
15	Malaysia	12	3.17	3.00	0.946	Lesotho	21	3.23	3.06	0.950
16	Tanzania	19	3.00	2.84	0.947	Egypt	13	3.19	3.03	0.949
17	China	1	2.94	2.78	0.946	India	4	2.81	2.67	0.949
18	Switzerland	8	2.84	2.69	0.946	Brazil	9	2.81	2.67	0.949
19	Russia	1	2.74	2.60	0.946	Saudi Arabia	1	2.81	2.67	0.949
20	Iran	7	2.47	2.33	0.946	Switzerland	8	2.78	2.64	0.949
21	Brazil	9	2.47	2.33	0.946	Russia	1	2.39	2.27	0.949
22	India	4	2.40	2.27	0.946	Iran	7	2.27	2.15	0.947
23	Argentina	6	1.99	1.89	0.945	Argentina	6	1.84	1.75	0.947
24	Equ. Guinea	24	1.41	1.34	0.945	Togo	22	1.55	1.47	0.947
	Total	181	100.00	100.00			184	100.00	100.00	
	Gini		0.280	0.317				0.270	0.304	

Power indices have been calculated using the method of generating functions using the program *ipgenf* on the website www.warwick.ac.uk/~ecaae

governments and the rest are elected to represent other countries which are arranged into constituencies around the candidate they voted for. The table shows, for the country of each director, the number of countries it represents, its voting weight in the Executive Board, its voting power, and the ratio of power to weight.

The USA dominates again but the inequality here, expressed by the Gini coefficient, inequality is less than it is in the Board of Governors because most directors cast the combined votes of their constituency members. The power ratios show that all directors lose power to the USA and have less power than weight. The USA has 25 % more voting power than weight. There is a slight reduction in inequality between 2006 and 2012, the weight of the USA falling from 17.09 to 16.79 %.

Inequality of voting power is slightly greater, falling slightly, that of the USA falling from 21.45 to 20.95 %. The Gini coefficient of the distribution of voting weight falls from 0.28 to 0.27 and that of the voting power indices from 0.317 to 0.304.

In so far as direct comparisons are meaningful, the results are similar to those for the Governors. We can make direct comparisons of power indices for the directly appointed directors, but they are not so straightforward for the elected directors because it is necessary to take account of the power distribution within the constituency. Some of the constituency directors can be thought of in the same way as the appointed directors because they dominate their constituencies, and therefore have the absolute power to cast the combined votes. They are—to use the language of the voting power literature—technically *dictators* within their constituency.[11] But others are elected and it is necessary to allow for the distribution of voting power within the constituency as well as the bloc vote cast by the elected director. We provide a fuller analysis of this feature of the Executive Board in Sect. 6 below. Here we simply treat the Executive Board as a single weighted voting body.

Table 3 shows the analysis for the 16 directors who represent a constituency with more than one member. The table shows, for the country of each director, for 2006 and 2012, its voting weight as a member of the Board of Governors, its voting weight as an executive director (the combined weight of all its constituency members) and the difference. The countries are ordered by the difference, which measures the gain in voting weight due to the constituency system. The table also includes the voting power of the country in its constituency assuming an election by simple majority vote [columns (3) and (6)].[12] Countries which are 'dictators' in their constituency have a voting power of 100 %.

The results show the countries that gain most in voting weight by the constituency system: Mexico (casting the votes of Venezuela, Spain and the Central American republics), Belgium (representing ten east European countries including Austria, Turkey, Hungary), Finland (representing the Nordic group), Tanzania (representing the Anglophone African group), and so on. In all constituencies where there is a member with over half the votes, and is a 'dictator', that member is always elected (Switzerland, Brazil, Italy, Canada, India). The gain in voting power is less in these cases simply because their voting power is high anyway. Where there is a member who is dominant in the constituency, such as the Netherlands (voting power index in the constituency 98.9 % in 2006, 89.1 in 2012), Belgium (voting power 68.0 and 53.0 %, respectively), Argentina (75 %) it is elected although not technically a 'dictator'. Other constituencies operate a more

[11] The constituencies are formed endogenously during the voting process: they have no objective status in the rules of the IMF. Members are free to leave and join another constituency by voting for another candidate in the biennial election of the board. Although voting patterns and therefore constituency membership are stable over time, migrations do occur. For example, Kazakhstan and Uzbekistan both changed constituency between 2006 and 2012, the former moving from the Belgian to the Swiss constituency, the latter from the Swiss to the Australasian constituency.

[12] Normalised Banzhaf index.

Table 3 Elected Executive Directors' votes

2006					2012				
	Weight (%)	ED wt. (%)	diff.	Const. VP (%)		Weight (%)	ED wt. (%)	diff	Const. VP (%)
	(1)	(2)		(3)		(4)	(5)		(6)
Mexico	1.2	4.27	3.07	33.3	Singapore	0.59	3.94	3.35	15.1
Belgium	2.13	5.13	3.00	68.0	Mexico	1.47	4.66	3.19	33.3
Finland	0.59	3.51	2.92	13.7	Lesotho	0.04	3.23	3.19	1.5
Tanzania	0.1	3	2.9	4.3	Belgium	1.86	4.98	3.12	53.0
Egypt	0.45	3.26	2.81	12.9	Egypt	0.4	3.19	2.79	12.9
Korea	0.76	3.33	2.57	13.3	Denmark	0.78	3.41	2.63	25.0
Malaysia	0.69	3.17	2.48	18.8	Netherland	2.08	4.53	2.45	89.1
Netherland	2.38	4.85	2.47	98.9	Australia	1.31	3.63	2.32	29.4
Iran	0.7	2.47	1.77	32.1	Iran	0.62	2.27	1.65	32.1
Equ. Guinea	0.03	1.41	1.38	1.8	Togo	0.06	1.55	1.49	3.7
Switzerland	1.6	2.84	1.24	100	Switzerland	1.4	2.78	1.38	100
Brazil	1.41	2.47	1.06	100	Italy	3.16	4.27	1.11	100
Argentina	0.98	1.99	1.01	75.0	Brazil	1.72	2.81	1.09	100
Italy	3.25	4.18	0.93	100	Canada	2.56	3.61	1.05	100
Canada	2.94	3.71	0.77	100	Argentina	0.87	1.84	0.97	75.0
India	1.92	2.4	0.48	100	India	2.34	2.81	0.47	100

Columns (1) and (4) are the country's weight share, (2) and (5) the combined constituency weight share of all countries in the constituency, the votes that the executive director casts, (3) and (6) voting power shares of the country within the constituency

open system of representation, with different countries providing the elected executive director (for example, the Nordic, Anglophone/Lusophone African and Francophone African constituencies). These conclusions are drawn from the observed voting weights and not from analysis of the operation of the constituencies in practice, which exist outside the formal rules of the IMF, and have adopt their own procedures.

The main result for the Executive Board is the same as for the Board of Governors: a strong tendency for weighted voting to enhance the voting power of the United States at the expense of all the other directors. The effect is not so great: here the ratio of voting power to voting weight for the USA is 1.248, showing that the USA gains a hidden extra share of voting power of almost 25 % more than its weight, compared with 45 % in the Board of Governors in 2012.

6 The Executive Board as a Representative Democratic Body

Executive directors have a dual role: on the one hand they are professional officers of the IMF who are permanently based in Washington, experts charged with designing and implementing policies that are supposed to be technically objective and politically neutral, and on the other they are either appointed or elected by member countries and therefore political representatives or delegates. The latter set of roles are our concern in this section: our focus is on the power relations between member countries.

Although the Articles prescribe a set of formal rules for electing directors, which do not mention constituencies at all, in practice the constituencies are a real force. There are no formally laid down rules governing the relationships between directors and their constituents that we can study. But in practice those member countries which do not have the right to appoint their own director are arranged into geographical groupings, whose members vote for the executive director who represents them. It is therefore natural to treat them as constituencies, since they are defined by the fact of the members voting for the executive director who casts their votes on their behalf.

Many of the constituencies have a powerful dominant member whose director is invariably elected—not least because (in five cases) his or her country has an absolute majority of the constituency votes—and so in effect these have become permanent board members. In these cases the other constituency members have no voting power in relation to the Executive Board. Two other constituencies have a practically dominant member who is not technically a dictator: those represented by Belgium and Netherlands. The other nine constituencies have no single dominant member and the chair rotates or changes otherwise.

In the discussion of the IMF it is customary to refer to the constituencies as if they operated just like any other in a representative democracy. Spokesmen for the IMF often refer to constituencies in these terms. Directors meet their constituencies at the annual IMF/World Bank meetings.

However, there appears to be an issue of democratic legitimacy when one reads in the authoritative work on the governance of the IMF: "When members belonging to a given constituency hold different views on a subject, the executive director can put differing views on record but cannot split his or her vote. The resolution of such conflicts is for each director to decide and any director remains free to record an abstention or an objection to a particular decision. The system has a tempering impact and evidence shows that the decisions that finally result may well be the best that could be taken under the circumstances" (Van Houtven 2002).

We can distinguish two types of constituencies in terms of their composition by types of countries that make them up. Seven are mixed industrial, middle income and developing or transitional countries and nine are developing countries. Many of them, especially the mixed groups, have a member with a very large weight, usually an industrial country, which is dominant within the group and whose representative is invariably elected. Some constituencies have different arrangements for selecting their director and the office rotates; this may be the case where there is no one member who is dominant in terms of weight, such as the Nordic-Baltic constituency and also the two African constituencies; alternatively there may be two or three relatively dominant members among whom the office rotates but excluding the smaller members, for example the Mexican-Venezuelan-Spanish group where there are three dominant members.

The Articles do contain one explicit provision for majority voting within constituencies: the procedure for a by-election for an executive director when there is a casual vacancy. The members of the relevant constituency elect the replacement director, by a simple majority of the votes cast, using eliminating ballots if necessary.[13] There has been at least one case where a constituency has actually elected its director by open voting. The Middle Eastern constituency in the IMF, which includes Egypt, Iraq, Kuwait and ten other Arab countries, has selected its executive director by open election between candidates from different countries. We therefore consider it is of interest and appropriate to investigate the voting power of the member countries using voting power analysis on the stylized model of representative democracy suggested by the constituency structure.

The first result of this analysis is that five members are formally 'dictators' within their constituency, all the other members are powerless. This applies to the constituencies of Italy, Canada, Switzerland, Brazil and India. Those countries which are rendered powerless, in the sense that their a priori voting power is zero, are referred to in the voting power and game theory literature as technically *dummies*. Uzbekistan was a 'dummy' in 2006 when it was a member of the constituency represented by Switzerland, but ceased to be so in 2012 when it had moved to the constituency represented by Australia; Kazakhstan became a 'dummy' when it moved from having voting power as a member of the constituency represented by Belgium in 2006 to that represented by Switzerland in 2012.

The fact that countries are 'dummies' when there is a 'dictator' is perhaps not a surprising finding. However, we have discovered that some countries have zero voting power although their constituencies do not have a dictator. This finding is illustrative of the value of the voting power approach because it is not obvious and could not have been discovered any other way. The countries are Estonia in 2006, and the five Central American countries, Costa Rica, El Salvador, Guatemala, Honduras and Nicaragua. That the latter five countries are dummies follows from the fact that their constituency has three large members, Spain, Mexico, Venezuela, any two of which are needed to form a majority, which implies that none of the other five members can ever be decisive. However, the finding that Estonia was a 'dummy' in the Nordic constituency is not at all obvious. It is a property of the voting weight the country receives by virtue of its quota. Note that Estonia has positive voting power by 2012 following the quota reforms.

Therefore, in 2006 there were in total 42 member countries (23 % of the membership), with zero voting power with respect to the Executive Board, in possession of some 4.19 % of the voting weight. The reforms made little difference to this: by 2012 the number had fallen to 41 with slightly more, 5.55 %, of the weight. These countries include some industrial countries but in the main they are developing countries. They are: Albania, Antigua and Barbuda, Azerbaijan, Bahamas, Bangladesh, Barbados, Belize, Bhutan, Colombia, Costa Rica, Dominica, Dominican Republic, Ecuador, El Salvador, Estonia (2006), Greece,

[13] Article XII, Sect. 3 (f), and By-Law 17.

Grenada, Guatemala, Guyana, Haiti, Honduras, Ireland, Jamaica, Kazakhstan (2012), Kyrgyz Republic, Malta, Nicaragua, Panama, Poland, Portugal, San Marino, Serbia, Sri Lanka, St. Kitts and Nevis, St. Lucia, St. Vincent and the Grenadines, Suriname, Tajikistan, Timor-Leste, Trinidad and Tobago, Turkmenistan, Uzbekistan (2006). The six 'dummies' with the largest weight are Poland, Portugal, Ireland, Greece, Colombia, Bangladesh.

We now analyse the voting power of every member by considering the Executive Board as a two-level representative body. Each member's voting power is the product of voting power in two voting bodies: its power with respect to decisions taken by simple majority voting among constituency members within the constituency, and the power of the constituency in the Executive Board under simple majority voting. A member's power index is obtained by multiplying together these two Penrose indices.[14]

It is of interest to use this technique to investigate which members gain and which lose power in the Executive Board as a result of the way the constituency system is assumed to work, compared with their power in the Board of Governors. This comparison assumes away the differences in competence of the two bodies and focuses only the structural effects of the constituency groupings on formal voting power. Obviously the members who have been shown to be powerless in their constituencies are losers. However, it is not clear that the countries that dominate their constituencies, including the 'dictators', necessarily gain since it depends on the power of their constituency.

Table 4 gives some results of this analysis. Only the results for the most powerful countries are presented. The power indices for the Board of Governors, from Table 1, are also presented as the basis of comparison with the indices for the two-level voting structure we have assumed. For each year the countries are ordered by their two-stage indirect voting power index. From these results we can infer that, in 2006, the countries that most benefited from the constituency system—that is with both a large indirect voting power index and with its indirect power index greater than its direct (the Board of Governors) power index—are the Netherlands (3.76 % compared with 2.18 %), Belgium (a very large increase: 3.69 % compared with 1.95 %), Italy, Canada, Switzerland. The same pattern was repeated in 2012.

Table 5 shows the biggest gainers and losers from the constituency system in terms of voting power. Here the countries are ordered by gain or loss, that is the difference between two-stage indirect power index for the Executive Board and the direct power index for the Board of Governors (labelled VP). The biggest gainers are the rich smaller European countries especially Belgium and the Netherlands. The biggest losers tend to be the countries that appoint their own Directors: the USA, Japan, Germany, UK, France. By 2012 China had become big enough for this effect to apply to it. The biggest losers also include 'dummy' countries Poland and Ireland.

[14] The absolute (that is, non-normalised) power indices, which are probabilities, are used for this calculation. The normalised indices are then computed.

Table 4 Voting power indices for the Executive Board as a democratic representative body (two-stage voting) (most powerful countries)

2006				2012			
	Weight (%)	2-stage VP (%)	BGVP (%)		Weight (%)	2-stage VP (%)	BGVP (%)
USA	17.09	17.57	24.49	USA	16.79	16.46	24.29
Japan	6.13	4.77	5.46	Japan	6.25	4.68	5.50
Germany	5.99	4.66	5.35	Germany	5.82	4.36	5.18
France	4.95	3.84	4.48	Netherland	2.08	3.33	1.90
UK	4.95	3.84	4.48	France	4.30	3.21	3.89
Netherland	2.38	3.76	2.18	UK	4.30	3.21	3.89
Belgium	2.13	3.69	1.95	Italy	3.17	3.19	2.88
Italy	3.25	3.25	2.97	Belgium	1.86	3.13	1.70
Canada	2.94	2.88	2.69	China	3.82	2.85	3.46
Saudi Arabia	3.22	2.50	2.94	Canada	2.56	2.70	2.34
China	2.94	2.28	2.69	India	2.35	2.10	2.14
Switzerland	1.60	2.20	1.46	Brazil	1.72	2.10	1.57
Russia	2.74	2.13	2.50	Saudi Arabia	2.81	2.09	2.56
Australia	1.50	2.04	1.37	Switzerland	1.41	2.07	1.28
Sweden	1.11	1.95	1.02	Russia	2.39	1.79	2.18
Brazil	1.41	1.91	1.29	Spain	1.63	1.74	1.49
India	1.92	1.86	1.76	Mexico	1.47	1.74	1.35
Spain	1.41	1.66	1.29	Venezuela	1.09	1.74	1.00
Venezuela	1.23	1.66	1.13	Sweden	0.98	1.47	0.90
Mexico	1.20	1.66	1.10	Korea	1.37	1.41	1.25
…	…	…	…	…	…	…	…

BGVP: Voting power in the Board of Governors (normalised Banzhaf index); 2-stage VP: normalised voting power index for idealised two-level voting in the Executive Board

Table 5 Voting power indices for the Executive Board as a democratic representative body (two-stage voting) (biggest gainers and losers)

2006				2012			
Gainers							
	2-stage VP (%)	BGVP (%)	Diff.		2-stage VP (%)	BGVP (%)	Diff.
Belgium	3.69	1.95	1.74	Netherland	3.33	1.90	1.43
Netherland	3.76	2.18	1.58	Belgium	3.13	1.70	1.42
Sweden	1.95	1.02	0.94	New Zealand	1.30	0.35	0.94
Indonesia	1.63	0.89	0.74	Switzerland	2.07	1.29	0.79
Switzerland	2.20	1.46	0.74	Venezuela	1.74	1.00	0.74
Australia	2.04	1.37	0.67	South Africa	1.30	0.71	0.59
Brazil	1.91	1.29	0.62	Indonesia	1.38	0.78	0.59
Kuwait	1.17	0.60	0.57	Sweden	1.47	0.90	0.57
Mexico	1.66	1.10	0.56	Brazil	2.10	1.57	0.52
Losers							
Austria	0.30	0.80	−0.50	Ireland	0.00	0.49	−0.49
Ukraine	0.00	0.59	−0.58	China	2.85	3.47	−0.62
Poland	0.00	0.59	−0.59	Poland	0.00	0.64	−0.64
France	3.84	4.48	−0.63	France	3.21	3.90	−0.68
UK	3.84	4.48	−0.63	UK	3.21	3.90	−0.68
Germany	4.66	5.35	−0.69	Germany	4.36	5.19	−0.83
Japan	4.77	5.46	−0.69	Japan	4.68	5.51	−0.83
USA	17.57	24.49	−6.92	USA	16.46	24.33	−7.87

BGVP: Voting power in the Board of Governors (normalised Banzhaf index); 2-stage VP: normalised voting power index for idealised two-level voting in the Executive Board

Table 6 Biggest gainers and losers in voting power 2006–2012

Gainers				Losers			
	VP2012	VP2006	Diff.		VP2012	VP2006	Diff.
Governors							
China	3.47	2.69	0.786	Netherland	1.90	2.18	−0.272
Korea	1.25	0.70	0.556	Russia	2.19	2.50	−0.316
India	2.14	1.76	0.386	Canada	2.34	2.69	−0.345
Brazil	1.57	1.29	0.283	Saudi Arabia	2.56	2.94	−0.376
Mexico	1.35	1.10	0.248	France	3.90	4.48	−0.577
Spain	1.49	1.29	0.200	UK	3.90	4.48	−0.577
Executive Board							
China	3.63	2.78	0.845	Netherland	4.30	4.59	−0.292
Singapore/Malaysia	3.75	3.00	0.743	Russia	2.27	2.60	−0.325
India	2.67	2.27	0.402	Saudi Arabia	2.67	3.05	−0.385
Mexico	4.43	4.05	0.377	USA	20.95	21.45	−0.505
Brazil	2.67	2.33	0.335	UK	4.09	4.69	−0.605
Australia/Korea	3.45	3.15	0.302	France	4.09	4.69	−0.605
Executive Board as a representative body (2-stage voting)							
Korea	1.41	0.54	0.863	Congo, Rep	0.14	0.67	−0.532
New Zealand	1.30	0.54	0.754	Belgium	3.13	3.69	−0.566
China	2.85	2.28	0.572	France	3.21	3.84	−0.630
Congo, DR	0.63	0.11	0.526	UK	3.21	3.84	−0.630
Singapore	0.92	0.50	0.420	Australia	1.31	2.04	−0.730
Malaysia	1.17	0.83	0.342	USA	16.46	17.57	−1.108

7 Conclusions

We have used the method of voting power analysis and power indices to analyse the voting system by which the IMF is governed and the recent reforms that have been made to it. We argue, and hopefully have demonstrated, that this approach provides valuable insights that help us better understand weighted voting systems. The results for the voting power implications of the recent reforms are summarized in Table 6.

We report three analyses of the reforms: first, their effect on the voting power relations in the Board of Governors, where all member countries have a voice; second, their implications for the distribution of voting power in the Executive Board among the 24 countries that are members of that body; and third, their implications for voting power with respect to representation on the Executive Board, where it is regarded formally as a delegate body using a two-stage voting procedure. The principal finding, from the first analysis, is that the voting power share of the United States is always substantially much more than its weight, while for all other members, their voting power shares are slightly lower than their weights. Not only is the allocation of voting weight very unfair from the point of view of an ideal of "One person, One vote", with many large developing countries and emerging markets seriously under-represented, but this bias is compounded by the inequality in the distribution of weights and the voting rules.

Table 6 reports the six biggest gainers and six biggest losers in terms of voting power shares for each of these three analyses. First, while the biggest gainers from the reforms in the Board of Governors are the emerging markets including China, Korea, India, Brazil and Mexico, the effects are quite small: for example the largest increase is that for China which is less than one percent of the total voting power. These increases are mainly at the expense of the voting power of some of the industrial countries including the biggest losers UK, France, Canada, and the Netherlands. All these effects are small which suggests that the reforms do not live up to some of the claims that have been made for them.

The second analysis is of the changes in the Executive Board. Again the biggest gainers are the emerging markets, China, India, Mexico, Brazil and the South–East Asian constituency represented either by Singapore or Malaysia. The biggest losers in voting power are the large industrial countries, this time including the USA.

The third analysis, of changes to voting power of member countries in relation to the Executive as a delegate body, shows more mixed results, with the biggest gainers being Korea, New Zealand and China, and the biggest losers the USA, Australia, the UK and France. All these effects are small however, and the overall conclusion must be that the reforms are insubstantial.

The results for the idealized two-level voting system we have assumed for the Executive Board and its constituencies suggest that such a system tends strongly to benefit the smaller developed European countries, notably Belgium and the Netherlands. Also, from this point of view, almost a quarter of all members, mostly small developing countries, are completely powerless.

These results point to a serious limitation in the democratic legitimacy of the governance of the institution. The recent quota reforms, while claimed as being a major step towards improving the voice and representation of the poor countries and emerging economies are nothing of the sort. The changes are very small and give no cause for celebration.

References

Brams, S. J., & Affuso, P. J. (1976). Power and size: a new paradox. *Theory and Decision, 7*(1–2), 29–56.

Buira, A. (2002). *A new voting structure for the IMF*. Washington: G24, www.g24.org/.

Buira, A. (Ed.). (2003). *Challenges to the World Bank and IMF: developing country perspectives*. London: Anthem Press.

Buira, A. (Ed.). (2005). *The IMF and World Bank at sixty*. London: Anthem Press.

Buira, A. (Ed.). (2006). *Reforming the governance of the IMF and the World Bank*. London: Anthem Press.

Coleman, J.S. (1971). Control of collectivities and the power of a collectivity to act. In: Lieberman, B. (Ed.), *Social choice*, New York, Gordon and Breach; reprinted in J.S. Coleman, 1986, *Individual interests and collective action,* Cambridge University Press.

Coleman, J. S. (1973). Loss of power. *American Sociological Review, 38*, 1–17.

Felsenthal, D. S., & Machover, M. (1998). *The measurement of voting power*. Cheltenham: Edward Elgar.

Leech, D. (2002a). Voting power in the governance of the IMF. *Annals of Operations Research, 109*, 373–395.

Leech, D. (2002b). Computation of power indices. Warwick Economic Research Papers, No. 644, July 2002.

Leech, D. (2002c). An empirical comparison of the performance of classical power indices. *Political Studies, 50*(1), 1–22.

Leech, D. (2003a). Computing power indices for large voting games. *Management Science, 49*(6), 831–838.

Leech, D. (2003b). The utility of voting power analysis. *European Union Politics, 4*(4), 479–486.

Leech, D. (2011). Power indices in large voting bodies. *Public Choice*, Online First, 10 Aug 2011.

Leech, D., & Leech, R. (2003). Website: "computer algorithms for voting power indices". www.warwcik.ac.uk/∼ecaae/.

Leech, D., & Leech, R. (2005a). Voting power implications of a unified European representation at the International Monetary Fund. Warwick Economic Research Papers, number 720, January 2005.

Leech, D., & Leech, R. (2005b). Voting power in the Bretton woods institutions. *Homo Oeconomicus, 22*(4), 1–23.

Leech, D., & Leech, R. (2006). Voting power and voting blocs. *Public Choice, 127*(3–4), 285–303.

Moggeridge, D. (Ed.) (1980). Collected works of John Maynard Keynes. Cambridge University Press, Cambridge (vol XXV).

Owen, G. (1995). *Game theory*. New York: Academic Press.

Van Houtven, L. (2002). *Governance of the IMF: decision making, institutional oversight, transparency and accountability*. IMF Pamphlet Series no. 53, IMF: Washington.

Woods, N. (2001). Making the IMF and the World Bank more accountable. *International Affairs, 77*(1), 83–100.

Woods, N. (2006). The globalisers: The IMF, the World Bank and their Borrowers, Cornell Studies in Money, Cornell University Press.

A Priori Voting Power and the US Electoral College

Nicholas R. Miller

1 Introduction

The President of the United States is elected, not by a direct national popular vote, but by an Electoral College system in which (in almost universal practice since the 1830s) separate state popular votes are aggregated by adding up electoral votes awarded on a winner-take-all basis to the plurality winner in each state. State electoral votes vary with population and at present range from 3 to 55. The Electoral College therefore generates the kind of weighted voting system that invites analysis using one of the several measures of *a priori* voting power. With such a measure, we can determine whether and how much the power of voters varies from state to state and how voting power would change under various alternatives to the existing Electoral College system.

With respect to the first question, directly contradictory claims are commonly expressed. Many commentators see a substantial small-state advantage in the existing system but others see a large-state advantage. Partly because the Electoral College is viewed by some as favoring small states and by others as favoring large states, it is commonly asserted that a constitutional amendment modifying or abolishing the Electoral College could never be ratified by the required 38 states. The so-called "National Popular Vote Plan" (an interstate compact among states with at least 270 electoral votes that would pledge to cast their electoral votes for the "national popular vote winner") has been proposed as a way to bypass the constitutional amendment process.

An earlier version of this chapter was published in *Homo Oeconomicus*: 26/3-4, 2009 (*Essays in Honor of Hannu Nurmi*, Vol. I, edited by Manfred J. Holler and Mika Widgrén).

N. R. Miller (✉)
Department of Political Science, University of Maryland Baltimore County (UMBC),
Baltimore, Maryland 21250, USA
e-mail: nmiller@umbc.edu

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_22,

The divergent assessments of bias in the Electoral College often arise from a failure by commentators to make two related distinctions. The first is the theoretical distinction between *voting weight* and *voting power*. The second is the practical distinction between *how electoral votes are apportioned* among the states (which determines their voting weights) and *how electoral votes are cast* by states (which influences their voting power).

These distinctions were clearly recognized many years ago by Luther Martin, a Maryland delegate to the convention that drafted the U.S. Constitution in 1787. Martin delivered a report on the work of the Constitutional Convention to the Maryland State Legislature in which he made the following argument (1787:198–199 all emphasis in original).

> [E]ven if the States who had the most inhabitants ought to have the greatest number of delegates [to the proposed House of Representatives], yet the *number of delegates* ought not to be in *exact proportion* to the *number of inhabitants* because the influence and power of those states whose delegates are numerous will be *greater*, when compared to the influence and power of the other States, than the *proportion* which the numbers of their delegates bear to each other; as, for instance, though Delaware has *one* delegate, and Virginia but *ten*, yet Virginia has *more* than *ten times* as much *power* and *influence* in the government as Delaware.[1]

Martin evidently assumed that each state delegation in the House would cast its votes as a bloc, so he counted up various voting combinations of states in order to support his claim. Martin's objection to apportioning seats proportionally to population correctly anticipated one of the fundamental propositions of modern voting power analysis—namely, that voting power may not be proportional to voting weight. This principle is most evident in the extreme case in which a single voter has a bare majority of the voting weight and therefore all the voting power, an example of which Martin provided earlier in his report (1787: 182).

Of course, Martin's expectation that state delegations in the House would cast bloc votes was not borne out. However, as noted at the outset, state electoral votes for President would soon be cast in blocs, and the U.S. Electoral College has subsequently been one of the principal institutions to which voting power analysis has been applied.

The mode of *apportioning* electoral votes is fixed in the Constitution: each state has electoral votes equal to its total representation in Congress, i.e., its House seats (apportioned on the basis of its population but with every state guaranteed at least one) plus two (for the Senators to which every state is entitled). Thus each state is guaranteed three electoral votes, and the apportionment reflects population only above this floor. The relative magnitude of the small-state advantage in electoral votes is determined by the ratio of the size of the Senate to the size of the House. While this ratio varied between about 0.19 and 0.29 during the nineteenth century,

[1] Cited by Riker (1986).

it has remained essentially constant at about 0.22 over the last 100 years.[2] Since 1912 the size of the House has been fixed at 435, since 1959 there have been 50 states, and since 1964 the 23rd Amendment has granted three electoral votes to the District of Columbia, so the total number of electoral votes at present is 538, with a bare majority of 270 votes required for election. Since 1964 a 269-269 electoral vote tie has been possible, so a Presidential election may be "thrown into the House of Representatives" (for lack of an electoral vote winner) even in the absence of third-party candidates winning electoral votes. In this event, the Constitution provides that the House of Representatives will choose between the tied candidates, with each state delegation casting one vote.

Additional Representatives (and electoral votes) beyond the floor of three are apportioned among the states on the basis of population. Since 1940, the "Hill-Huntington" apportionment formula (a divisor method also known as the "Method of Equal Proportions") has been used for this purpose, even though this method appears to have a slight small-state bias (Balinski and Young 1982). Figure 1a shows the apportionment of House seats following the 2000 census in relation to the population of each state. Evidently approximate proportionality is achieved but, because apportionment must be in whole numbers, apportionment cannot be perfect. This *Whole Number Effect* is most conspicuous among small states, as is highlighted in Fig. 1b. Figures 2a and 2b show the present apportionment of electoral votes in relation to population. It is evident that the small-state advantage resulting from the three electoral vote floor more than outweighs the capriciously unfavorable way some small states (in particular Montana, the largest state with only one House seat) are awarded House seats.

The manner of selecting electors (and thereby the manner of casting electoral votes) is not fixed in the Constitution. Rather the Constitution empowers the legislature of each state to decide how to do this. In early years, the manner of selecting electors was subject to regular manipulation by politicians seeking state and (especially) party advantage (most notably in advance of the bitterly contested 1796 and 1800 elections). But since about 1836, with only few exceptions, electors in each state have been popularly elected on a "general ticket" and therefore have cast their electoral votes on the winner-take-all basis noted at the outset. Each party in each state nominates a slate of elector candidates, equal in number to the state's electoral votes and pledged to vote for the party's Presidential and Vice-Presidential candidates; voters vote for one or other slate (not individual electors) and the slate that wins the most votes is elected and casts its bloc of electoral votes as pledged. By standard voting power calculations (and as anticipated by Martin), this winner-take-all practice produces a large-state advantage that in some measure counterbalances the small-state advantage in the apportionment of electoral votes.

---

[2] However, Congress has the power to change the size of the House without a constitutional amendment, so in principle it can increase or decrease the small-state advantage in electoral vote apportionment by decreasing or increasing the size of the House. For an analysis of how House size can influence the outcome of Presidential elections, see Neubauer et al. (2003).

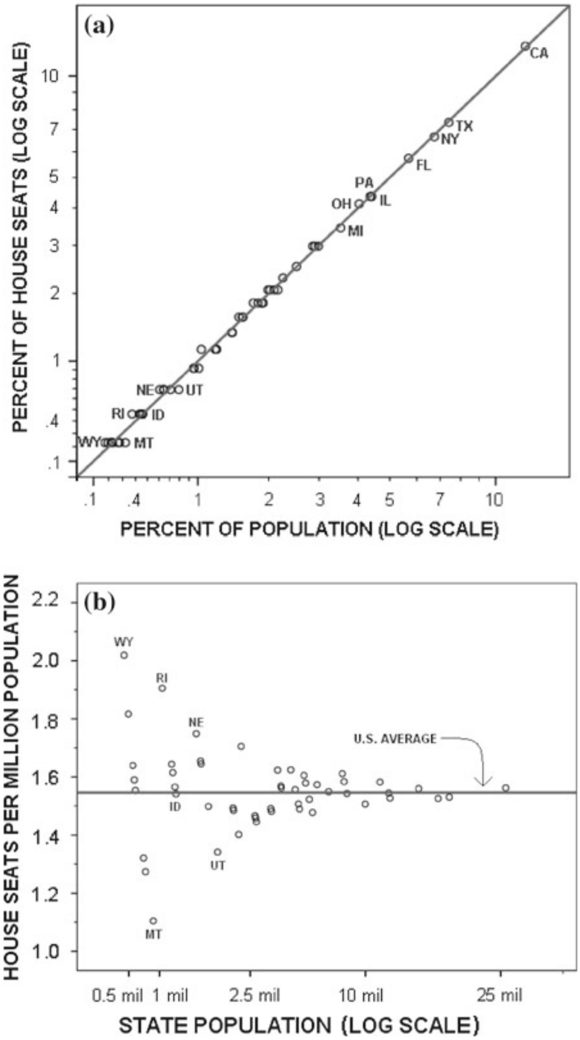
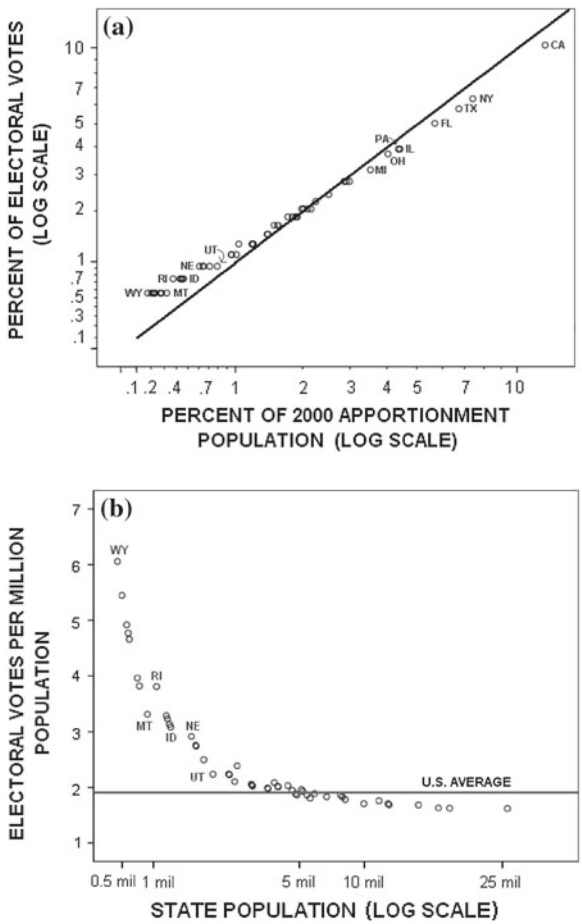


Fig. 1 **a** Apportionment of house seats (2000 census) by state population. **b** House seats per million by state population

2 Banzhaf Voting Power

A measure of *a priori* voting power takes account of the fundamentals of a voting rule but nothing else. Thus the following analysis takes account only of the 2000 population of each state and the District of Columbia, the apportionment of electoral votes based on that population profile, and the requirement that a Presidential candidate receive 270 electoral votes to be elected. It does not take account of other demographic factors, historical voting patterns, differing turnout rates, relative party strength, survey or polling data, etc. This indicates the sense in

Fig. 2 **a** Apportionment of electoral votes by state population. **b** Electoral votes per million by state population



which *a priori* voting power analysis is conducted behind a "veil of ignorance" and is blind to empirical contingencies.

In their authoritative treatise on the measurement of voting power, Felsenthal and Machover (1998) conclude that the appropriate measure of *a priori* voting power in typical voting situations, including the Electoral College, is the *absolute Banzhaf* (or *Penrose*) *measure* (Penrose 1946; Banzhaf III 1968). Like other voting power measures, the Banzhaf setup is based on votes and outcomes that are both binary in nature, as in a two-candidate election. The Banzhaf measure is defined as follows.

Given n voters, there are 2^{n-1} *bipartitions* (i.e., complementary pairs of subsets) of voters (including the pair consisting of the set of all voters and the empty set). A voter (e.g., a state) is *critical* in a bipartition if the set to which the voter belongs is *winning* (e.g., a set of states controlling 270 electoral votes) but would not be winning if the voter belonged to the complementary set. A voter's *Banzhaf score* is the total number of bipartitions in which the voter is critical. A voter's *absolute Banzhaf voting power* is the voter's Banzhaf score divided by the number of bipartitions.

While this ratio may seem *ad hoc* and without theoretical justification, it has an intuitive and coherent rationale in terms of probability. If we know nothing about a voting situation other than its formal rules, our *a priori* expectation is that everyone votes randomly, i.e., as if independently flipping fair coins. In such a *random voting* (or *Bernoulli*) *model*, each bipartition of voters into complementary sets, in which everyone in one set votes for one candidate and everyone in the other set for the other candidate, has equal probability of occurring. Therefore, *a voter's absolute Banzhaf voting power is the probability that the voter's vote is decisive*, i.e., determines the outcome, in what we may call a *Bernoulli election*.

In a simple one-person, one-vote majority-rule system with an odd number n of voters, the *a priori* voting power of a voter is the probability that the election is otherwise tied. If the number of voters n is even, the voter's voting power is one half the probability that the vote is otherwise within one vote of a tie. Provided that n is greater than about 25, this probability is very well approximated by the expression $\sqrt{2}/n$. This expression implies that, in a simple majority rule situation, *individual voting power* is inversely proportional, not to the number of voters, but to the *square root of the number of voters*. We refer to the $\sqrt{2}$/n formula as the *Inverse Square Root Rule* for simple majority rule voting. Given simple-majority rule *Bernoulli elections* with n voters, the expected vote for each candidate is 50 %, the probability that each candidate wins is 0.5, and the standard deviation of either candidate's absolute vote (over repeated elections) is approximately $0.5\sqrt{n}$.

Calculating Banzhaf voting power values in voting situations in which voters have unequal weights is considerably more burdensome. Direct enumeration by the Banzhaf formula informally sketched out above is feasible (even using a computer) only if the number of voters does not exceed about 25. It is possible to make exact calculations (using a computer) for up to about 200 voters by using so-called generating functions. Dennis Leech and Robert Leech have created a website for making voting power calculations using these and other methods.[3] I have used this website, together with the Inverse Square Root Rule, to make all the direct calculations reported below.

3 Voting Power in the Existing Electoral College

The *a priori* Banzhaf voting power of states in the current Electoral College is shown in Table 1. (Since equal electoral votes imply equal voting power, states need not be individually listed.) For the moment, I ignore the fact that Maine and

[3] *Computer Algorithms for Voting Power Analysis* (http://www.warwick.ac.uk/~ecaae/). The calculations in this chapter used its *ipgenf* algorithm.

Table 1 *A priori* state voting power in the current electoral college

EV	N	Absolute Banzhaf
3	8	0.022730
4	5	0.030312
5	5	0.037900
6	3	0.045493
7	4	0.053094
8	2	0.060704
9	3	0.068324
10	4	0.075955
11	4	0.083599
12	1	0.091257
13	1	0.098930
15	3	0.114328
17	1	0.129805
20	1	0.153194
21	2	0.161043
27	1	0.208805
31	1	0.241422
34	1	0.266331
55	1	0.475036
538	51	4.166201

EV: Number of Electoral Votes
N: Number of States
Absolute Banzhaf: Absolute Banzhaf voting power
Calculated by *ipgenf* at http://www.warwick.ac/~ecaae/

Nebraska actually award their electoral votes in the manner of the "Modified District System" discussed below. Remember that a state's Banzhaf voting power is the probability that its block of electoral votes is decisive in a Bernoulli election. Thus, with 55 electoral votes, California's Banzhaf voting power of about 0.475 means that, if we repeatedly flip fair coins to determine how each state other than California casts its electoral votes, about 47.5 % of the time neither candidate would have 270 electoral votes before California casts its votes and therefore either (a) California's bloc of 55 votes would determine the winner or (b) the leading candidate would have exactly 269 electoral votes and California's 55 votes would either elect that leading candidate or create a tie. (In the latter event, the Banzhaf measure in effect awards California "half credit.")

Figure 3 shows each state's share of voting power in relation to its share of electoral votes. Only California has a noticeably larger share of voting power than of electoral votes and, even for this mega-state, voting power only slightly exceeds voting weight. This is a manifestation of the *Penrose Limit Theorem*, which states that voting power tends to become proportional to voting weight as the number of

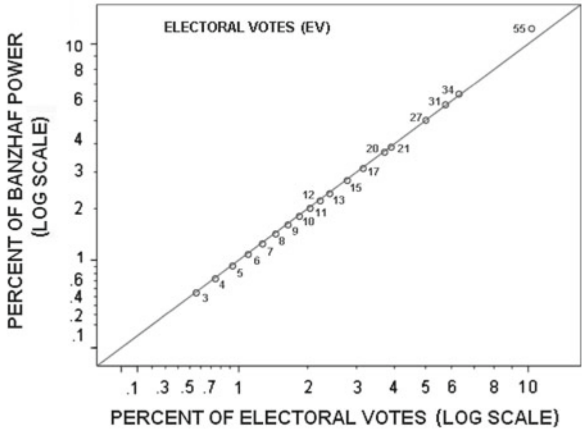


Fig. 3 Banzhaf voting power of states by electoral votes

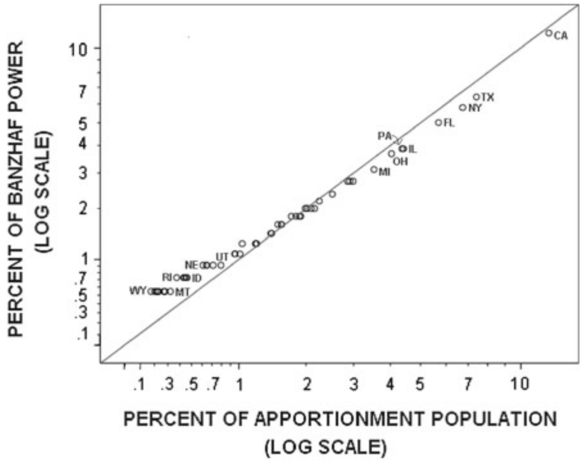


Fig. 4 Banzhaf voting power of states by population

voters increases, provided that the distribution of voting weights is not "too unequal."[4]

Figure 4 shows each state's share of voting power in relation to its share of the U.S. population. The small-state apportionment advantage still shows up quite prominently, and even California's noticeable advantage with respect to voting

[4] This "theorem" is actually a conjecture that has been proved in important special cases and supported by a wider range of simulations; see Lindner and Machover (2004) and Chang et al. (2006). The number of voters need not be very large in order for the theorem statement to be true to good approximation. Indeed, given the provisional apportionment of 65 House seats among only 13 states that was the focus of Luther Martin's objections, the Penrose Limit Theorem held to reasonable approximation (Virginia's advantage then was roughly comparable to California's today), so Martin's complaint about the disproportionate voting power of large states, while theoretically insightful, was in the circumstances largely off-the-mark (but also see Footnote 8).

power does not fully compensate for its disadvantage with respect to apportionment (though California does better than all intermediate-sized states).

But the 51-state Electoral College weighted voting system depicted in Fig. 3 is largely a chimera, since states are not voters but merely geographical units within which popular votes are aggregated. A U.S. Presidential election really is a *two-tier* voting system, in which the casting of electoral votes is determined by the popular vote within each state. So we now turn to the power of individual voters under the Electoral College system.

One distinct advantage of the absolute Banzhaf power measure is its applicability to two-tier voting systems such as the Electoral College. The voting power of an individual voter depends on both his voting power in the simple majority election within the voter's state and the voting power of that state in the Electoral College itself. Since both voting power values can be interpreted as probabilities, they can be multiplied together to get the voter's overall two-tier voting power. That is to say, the *a priori* voting power of an individual voter in the Electoral College system (as it works in practice) is the probability that the voter casts a decisive vote in his state (given by the Inverse Square Root Rule) multiplied by the probability that the bloc of votes cast by the voter's state is decisive in the Electoral College (given by the calculations displayed in Table 1) or, as we may say informally, the probability of "double decisiveness."

Putting the Inverse Square Root Rule and the Penrose Limit Theorem together (and referring to the units in the second tier generically as "districts"), we can derive the following expectations pertaining to two-tier voting systems. First, if the voting weight of districts is proportional to the number of voters in each, individual two-tier voting power *increases* proportionately with the square root of the number of voters in the voter's district. We call this the *Banzhaf Effect*. Second, if the voting weight of districts is equal (regardless of the number of voters in each), individual two-tier voting power *decreases* proportionately with the square root of the population of the voter's district. We call this the *Inverse Banzhaf Effect*. Given the preceding considerations, we can anticipate the approximate results of Banzhaf calculations of individual two-tier voting power under the Electoral College to be as follows.

1. Individual voting power within each state is inversely proportional to the square root of the number of voters in the state (due to the Inverse Square Root Rule).
2. As shown in Chart 3, state voting power in the Electoral College is approximately proportional to its voting weight, i.e., its number of electoral votes (due to the Penrose Limit Theorem).
3. As shown in Chart 2, the voting weight of states in turn is approximately (apart from the small-state apportionment advantage) proportional to population (and therefore to the number of voters).
4. As shown in Chart 4 and putting (2) and (3) together, state voting power is approximately proportional to population.
5. So putting together (1) and (4), individual *a priori* two-tier voting power is approximately proportional to the square root of the number of voters in a state.

Table 2 *A priori* individual voting power in selected states

State	Elect size	IND VP	EV	State VP	IND 2-T VP	REL VP
MT	392,640	0.00127334	3	0.022730	0.00002894	1.000000
UT	970,074	0.00081010	5	0.037900	0.00003070	1.060803
DE	340,488	0.00136738	3	0.022730	0.00003108	1.073857
NH	537,107	0.00108870	4	0.030312	0.00003300	1.140203
OK	1,500,107	0.00065145	7	0.053094	0.00003459	1.195039
AK	272,771	0.00152771	3	0.022730	0.00003472	1.199770
WS	2,329,521	0.00052277	10	0.075955	0.00003971	1.371895
CO	1,870,085	0.00058346	9	0.068324	0.00003986	1.377338
MD	2,302,057	0.00052587	10	0.075955	0.00003994	1.380054
MA	2,756,442	0.00048058	12	0.091257	0.00004386	1.515269
NC	3,498,990	0.00042655	15	0.114328	0.00004877	1.684919
MI	4,317,893	0.00038398	17	0.129805	0.00004984	1.722080
OH	4,933,195	0.00035923	20	0.153194	0.00005503	1.901409
IL	5,394,875	0.00034352	21	0.161043	0.00005532	1.911389
PA	5,334,862	0.00034544	21	0.161043	0.00005563	1.922110
FL	6,951,810	0.00030262	27	0.208805	0.00006319	2.183181
NY	8,242,552	0.00027791	31	0.241422	0.00006709	2.318163
TX	9,066,167	0.00026499	34	0.266331	0.00007057	2.438416
CA	14,715,957	0.00020799	55	0.475036	0.00009880	3.413738
US	122,294,000	0.00007215	538	–	0.00007215	2.492845

Elect size: Size of Electorate
[2000 Population × 0.4337, where 0.4337 = 2004 Total Presidential Vote/2000 US Population]
IND VP: Individual Absolute Banzhaf Voting Power within State
[by the Inverse Square Root Rule]
State VP: State Absolute Banzhaf Voting Power (from Table 1)
IND 2-T VP: Individual Banzhaf Voting Power in Two-Tier System [= IND VP × State VP]
REL VP: Relative Individual 2-T Voting Power (rescaled so that minimum [Montana] = 1)

However, this large-state advantage is counterbalanced in some degree by the small-state apportionment advantage.

In his pioneering analysis of voting power in the Electoral College (based on the 1960 Census), Banzhaf III (1968) reached the following conclusion.

> [A] voter in New York State has 3.312 times the voting power of a citizen in another part of the country…. Such a disparity in favor of the citizens of New York and other large states also repudiates the often voiced view that the inequalities in the present system favor the residents of the less populous states.[5]

[5] Specifically, Banzhaf found that voters in New York (the largest state at the time of the 1960 census) had 3.312 times the voting power of voters in the District of Columbia; they had 2.973 times the voting power of voters in the least favored state (Maine). The maximum disparity resulted from the stipulation in the 23rd Amendment that the District cannot have more electoral votes than the least populous state. In the 1960 census, the District not only had a population larger than every state with 3 electoral votes but also larger than several states with 4 electoral votes. The District today has a smaller population than every state except Wyoming.

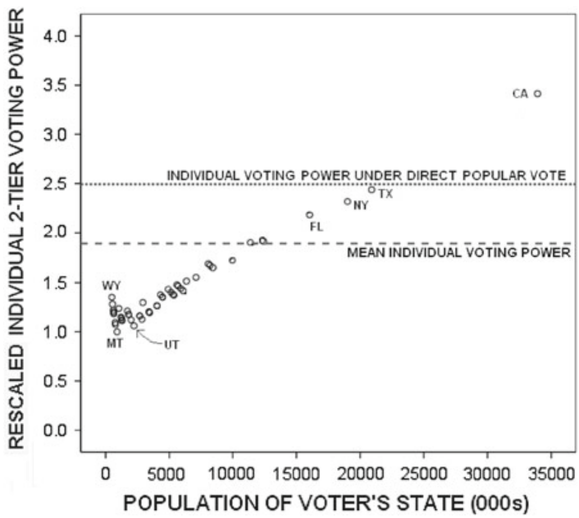


Fig. 5 Individual voting power by state population under the existing apportionment of electoral votes

Table 2 shows how *a priori* individual voting power under the existing Electoral College (based on the 2000 Census) varies across selected states. The calculations that underlie this table and the subsequent charts assume that the number of voters in each state is a fixed percent of its population in the 2000 Census.[6] The last column shows individual two-tier voting power rescaled in the manner suggested in the Banzhaf quotation above, i.e., so that the individual two-tier voting power of voters in the least favored state (Montana, the largest state with only three electoral votes) is 1.0000 and other power values are multiples of this. The last row shows individual voting power, under the same assumptions about electorate size, given direct (single-tier) popular election of the President.

Figure 5 shows rescaled individual two-tier voting power of voters in all 50 states plus the District of Columbia, the mean voting power of all voters in the existing Electoral College, and individual (single-tier) voting power given a direct popular election. Note that the latter is substantially greater than mean individual voting power under the Electoral College—indeed, it is greater than individual voting power in every state except California. So by the criterion of maximizing individual *a priori* voting power (which is hardly the only relevant criterion), only voters in California would have reason to object to replacing the

[6] This was taken to be 43.37 %, which is equal to the total popular vote for President in 2004 (122,294,000) as a percent of the U.S. population in 2000. *A priori*, we have no reason to expect that the percent of the population that is eligible to vote, or of eligible voters who actually do vote, varies by state (though, empirically and *a posteriori*, we know there is considerable variation in both respects). Using a different (fixed) percent of the population to determine the number of voters in each state would (slightly) affect the following estimates of absolute individual voting power but not comparisons across states or Electoral College variants.

Electoral College by a direct popular vote. On average, individual voting power would be about 1.35 times greater given a direct popular vote than under the existing Electoral College.

We noted earlier that a 269-269 electoral vote tie is a possibility. The preceding Banzhaf voting power calculations take account of the possibility of such a tie, but they do not take account of what happens in the event of such a tie as consequence of other provisions in the U.S. Constitution. What happens is that the election is "thrown into the House of Representatives," whose members choose between the two tied candidates, but with the very important proviso that each state delegation casts one vote. If we assume that voters in each state in effect vote for a slate of House members (just as they vote a slate of electors), individual voting power in this contingent runoff election is equal to what it would be if electoral votes were apportioned equally among the states (as discussed in the next section and depicted in Fig. 9). A fully comprehensive assessment of individual voting power under the existing Electoral College would take account of the probability of such a contingent indirect election. The one additional calculation needed is the probability that the Electoral College produces a 269-269 vote tie in a Bernoulli election, which is approximately 0.007708.[7]

Clearly the effect of taking account of this contingent procedure in which states have equal weight is to reduce the power of voters in large states and increase the power of voters in small states. However, the District of Columbia does not participate in this contingent election (since it has no Representative), so the power of voters in DC is reduced by the probability of an electoral vote tie. Only the power of California voters is noticeably (but not greatly) reduced, while the power of voters in states with up to about 7 electoral votes is noticeably (but not greatly) increased, when we take account of House runoffs.

More realistically, second-tier voting by state delegations in the event of an Electoral College tie is not (like states voting in the Electoral College) a chimera. Since Representatives would not have been elected by first-tier voters on the basis of their prospective vote for President, and state delegations would typically be internally divided, the more realistic conception is that, in the event of an electoral vote tie, ordinary voters drop out of the picture and individual House members vote in the first tier, their votes are aggregated within state delegations, and each state cast its single "electoral vote" accordingly in the second tier. From this more realistic point of view, the small-state advantage in the second-tier is substantially diluted by the Inverse Banzhaf Effect within delegations. For example, the second-tier voting power of each state delegation is 0.112275; the single Delaware Representative is always decisive within his delegation and thus has individual two-tier voting power

[7] This probability can be derived from other values calculated and displayed by the *ipgenf* algorithm of *Computer Algorithms for Voting Power Analysis*.

of 0.112275; each member of the 15-member Michigan delegation has first-tier voting power 0.209473 and therefore individual two-tier voting power of 0.023519—a bit over $^1/_5$ (not $^1/_{15}$) that of the Delaware member.[8]

4 Voting Power and the Apportionment of Electoral Votes

We next consider two types of variants of the existing Electoral College system. Variants of the first type retain the winner-take-all practice for casting electoral votes but employ different formulas for apportioning electoral votes among states. Variants of the second type retain the existing apportionment of electoral votes among the states but change the winner-take-all practice for casting electoral votes (or, in one case, adds "national" electoral votes).

Variants of the first type include the following:

(a) apportion electoral votes in whole numbers entirely on the basis of population (e.g., on the basis of House seats only);
(b) apportion electoral votes fractionally to be precisely proportional to population;
(c) apportion electoral votes fractionally to be precisely proportional to population but then add back the two electoral votes based on Senate representation; and
(d) apportion electoral votes equally among the states (in the manner of House voting on tied candidates), so that the winning candidate is the one who carries the most states.

Figure 6 shows individual voting power with electoral votes apportioned by House seats only. At first blush it may not look much different from Fig. 5, but it is important to take careful note of the vertical scale. Removing the small-state apportionment advantage in this way has the consequence of making voting power in the most favored state of California about ten times (rather than about three times) greater than that in Montana. Since apportionment is still in whole numbers, capricious inequalities remain among states with quite similar small populations.

Figure 7 shows individual voting power with electoral votes apportioned precisely by population (so states have fractional electoral votes). Since we have removed both the small-state apportionment advantage and the capricious effects of apportionment into whole numbers, the Banzhaf Effect as it pertains to

[8] Had Luther Martin's concern been the two-tier voting power of *individual members* of the House (rather than the voting power of *state delegations*) under the assumption of bloc voting by state delegations, his complaint that states should not have representation proportional to population would have been strongly supported by the theory of voting power measurement, because large-state members benefit from the (direct) Banzhaf Effect. Using the same example, under the Martin setup the Delaware and Michigan delegations have second-tier voting power of 0.008314 and 0.125606 respectively, so their members have two-tier voting power of 0.008314 and 0.026311 respectively, the latter being more than three times greater than the former.

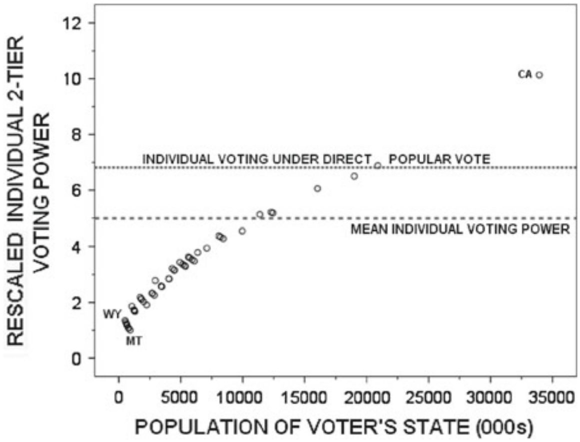


Fig. 6 Individual voting power by state population with electoral votes based on house seats only

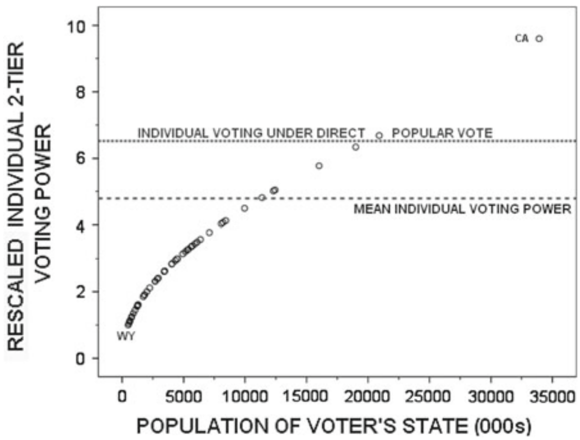


Fig. 7 Individual voting power by state population with electoral votes precisely proportional to population

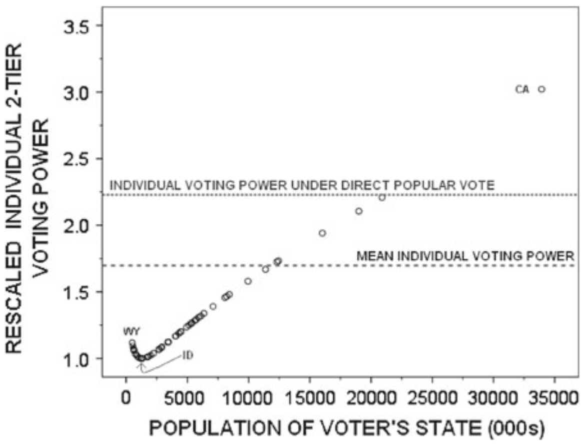


Fig. 8 Individual voting power by state population with electoral votes precisely proportional to population plus two

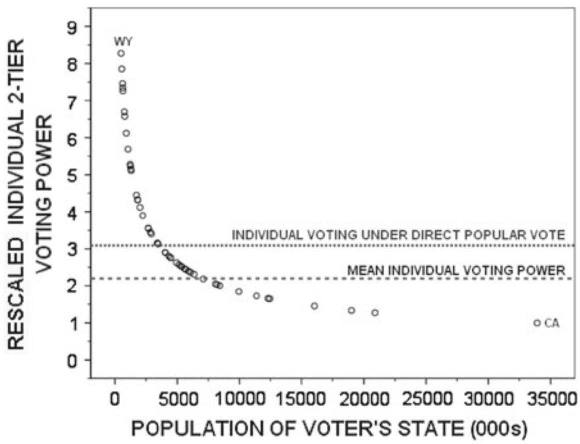


Fig. 9 Individual voting power by state population when states have equal electoral votes

individual first-tier voting power is essentially all that matters, and individual voting power increases smoothly and almost perfectly with the square root of state population. Voting power in California remains about ten times greater than voting power in the least favored state; however that least favored state is now Wyoming, as it is the smallest state and therefore has the least (fractional) electoral vote weight.

Figure 8 shows individual voting power with House electoral votes apportioned precisely by population but with the two Senatorial electoral votes added back in. Individual voting power continues to vary smoothly with population but in a non-monotonic fashion, as the relationship takes on a hockey-stick shape. The voting power advantage of California voters falls back again to about three times that of the least favored state. As the population of a voter's state increases, the small-state apportionment advantage diminishes but at a declining rate, while the large-state voting power advantage due to Banzhaf Effect increases. Idaho happens to be at the point on the population scale where these effects balance out, and it is therefore the least favored state.

Figure 9 shows individual voting power when all states have equal voting weight and is in a sense the inverse of Fig. 7. Since all states have equal second-tier voting power, individual voting power varies only with respect to first-tier voting power, and therefore smoothly reflects the Inverse Banzhaf Effect.

We may ask whether it is possible to apportion electoral votes among the states so that, even while retaining the winner-take-all practice, individual two-tier voting power is equalized across all states. One obvious but constitutionally impermissible possibility is to redraw state boundaries so that all states have the same number of voters (and electoral votes). Equalizing state populations in this way to create a system of *uniform apportionment* that not only equalizes individual voting power across states, but also increases mean individual voting power, relative to that under any type of apportionment based on actual and unequal state populations (Kolpin 2003). However, even while increased as well as equalized,

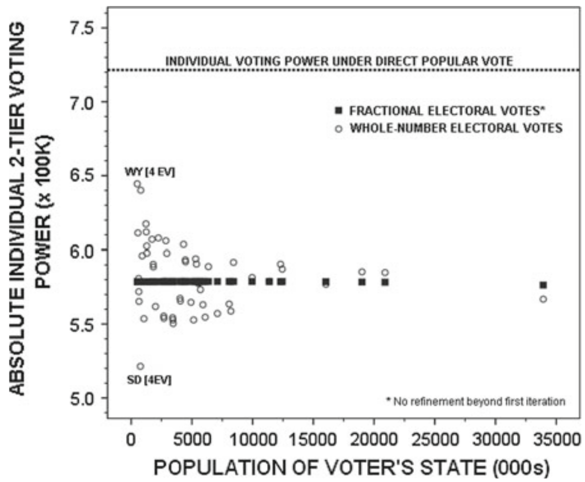


Fig. 10 Individual voting power by state population under penrose apportionment

individual voting power still falls below the (equal) individual voting power under direct popular vote. So the fact that mean individual voting power under the Electoral College falls below that under direct popular vote is not due mainly to the fact that states are unequal in population and cast their unequal electoral votes on a winner-take-all basis; rather it is evidently intrinsic to any strictly two-tier system.

Given that state boundaries are immutable, can we apportion electoral votes so that (without changing state populations and preserving the winner-take-all practice) the voting power of individuals is equalized across all states? Individual voting power can be equalized (to a high degree of perfection) by apportioning electoral votes so that state voting power is proportional to the square root of state population. This entails using what Felsenthal and Machover (1998, p 66) call the *Penrose Square Root Rule*. Such *Penrose apportionment* can be tricky, because what must be made proportional to population is not electoral votes (which is what we directly apportion) but state voting power (which is a consequence of the whole profile of electoral vote apportionment). However, in the case of the Electoral College we can immediately come up with an excellent approximation, which can be refined as desired. This is because, as we saw earlier, each state's share of voting power in the Electoral College is close to its voting weight, because $n = 51$ is large enough, and the distribution of state populations is equal enough, for the Penrose Limit Theorem to hold to very good approximation. So simply apportioning electoral votes to be precisely proportional (by allowing fractional electoral weight) to the square root of the population (or number of voters) in each state, we achieve almost perfect equality of voting power (call this *pure Penrose apportionment*); further refinement seems unnecessary, especially as electoral votes probably must be apportioned into whole numbers anyway. But even if we must be content with *whole-number Penrose apportionment*, we can make individual voting power much more equal than it is now or would be under any of the Electoral College variants examined here, other than the National Bonus Plan with a large bonus. Once again the whole-number effect capriciously advantages or disadvantages small states much more than large states, as is shown in Fig. 10. The chart once again makes clear that equalizing individual voting power is not the same as maximizing it (as under direct popular vote).

5 Voting Power and the Casting of Electoral Votes

We now consider the second type of variant of the existing Electoral College system. These variants retain the existing apportionment of electoral votes but employ rules other than winner-take-all for the casting of electoral votes.[9] Such variants include the following.

The Pure District Plan. Each state is divided into as many equally populated single-member districts as it has electors, and one elector is elected from each district. In effect, each party ticket earns one electoral vote for each district it wins.[10]

The Modified District Plan. Electors apportioned to a state on the basis of House seats are elected from the same equally populated Congressional Districts as the House members; the two additional electors apportioned to each state on the basis of their two Senate seats are (like the Senators) elected at-large. In effect, a

[9] Many of these Electoral College variants have actually been proposed as constitutional amendments, while few if any amendments have proposed changes in the apportionment of electoral votes. For a review of proposed constitutional amendments pertaining to the Electoral College, see Peirce and Longley (1981), especially chapter 6 and Appendix L. However, provided the position of Presidential elector is retained, each state legislature is free to change its mode of casting electoral votes (or, more directly, its mode of selecting Presidential electors) and, as previously noted, Maine and Nebraska actually depart from the winner-take-all arrangement at the present time.

[10] Evidently most members of the Constitutional Convention expected that electors would be popularly elected in this manner. However, their Constitution left this matter up to individual states legislatures. Under the original Electoral College system, each elector cast two undifferentiated votes for President. The candidate with the most votes became President (provided he received votes from a majority of electors) and the runner-up became Vice-President. After the first two contested Presidential elections in 1796 and 1800, it was clear that this system could not accommodate elections in which two parties each ran a ticket with both a Presidential and Vice-Presidential nominee. Following the election of 1800, there was considerable consensus to change the manner of casting electoral votes so that each elector would cast one designated vote for President and one designated vote for Vice President, and this was accomplished by the Twelfth Amendment. Though early drafts included the requirement that electors be popularly elected in the manner of the Pure District Plan, this provision was ultimately dropped from the amendment; see Kuroda (1994).

party ticket earns one electoral vote for each Congressional District it wins and two electoral votes for each state it wins.[11]

The Pure Proportional Plan. The electoral votes of each state are cast (fractionally) for party tickets in precise proportion to their state popular vote totals.[12]

The Whole Number Proportional Plan. The electoral votes of each state are cast in whole numbers for party tickets on the basis of an apportionment formula applied to the state popular vote.[13]

The National Bonus Plan. Existing electoral votes are apportioned and cast as at present, but the national popular vote winner earns an additional electoral vote bonus of some magnitude.[14]

Voting power calculations for the Pure District Plan can be made in just the same way as our previous results. Calculations for the other plans require somewhat different modes of analysis. In particular, those for the National Bonus and Modified District Plans present formidable difficulties, because each voter casts a single vote that counts in two ways.

Under the Pure District Plan, all voters in the same state have equal first-tier voting power, which can be calculated by the Inverse Square Root Rule, with n equal to the number of voters in the state divided by its number of electoral votes. Since the second-tier voting is also unweighted, the second-tier voting power of each district can also be calculated by the Inverse Square Root Rule with $n = 538$. Figure 11 shows individual voting power by state under the Pure District Plan. Inequalities come about entirely because of apportionment effects—in

[11] This is the system used at present by Maine (since 1972) and Nebraska (since 1992). The 2008 election for the first time produced a split electoral vote in Nebraska, where Obama carried one Congressional District; the Republican-dominated legislature may now switch state law back to winner-take-all. A proposed constitutional amendment (the Mundt-Coudert Plan) in the 1950s would have mandated the Modified District Plan for all states.

[12] A proposed constitutional amendment (the Lodge-Gossett Plan) along these lines was seriously considered in Congress in the late 1940s and 1950s. Since fractional electoral votes would be cast, the position of Presidential elector would necessarily be abolished, so this change can be effected only by constitutional amendment. Since minor candidates would presumably win (fractional) electoral votes, it becomes more likely that neither major candidate would win a majority of the electoral votes, so such an amendment would also have to specify what would happen in this event. (The Lodge-Gossett Plan would have elected the electoral-vote plurality winner, unless that candidate failed to receive at least 40 % of the electoral votes, in which case Congress voting by joint ballot would choose between the top two candidates ranked by electoral votes.).

[13] Since electoral votes would still be cast in whole numbers, the position of elector can be retained, and a state may use this formula unilaterally. Indeed, such a system was proposed in Colorado as initiative Proposition 36 in 2004. Since third candidates would be likely to win a few electoral votes (especially in large states), this system, if widely adopted, would throw more elections into the House.

[14] The principal purpose of such a plan is evidently to reduce the probability of an “election inversion” (Miller 2012) of the sort that occurred in 2000. The larger the national bonus, the more this probability is reduced. A bonus of 102 electoral votes has been most commonly discussed. (It would make sense, however, to make the bonus an odd number so as to preclude electoral vote ties.)

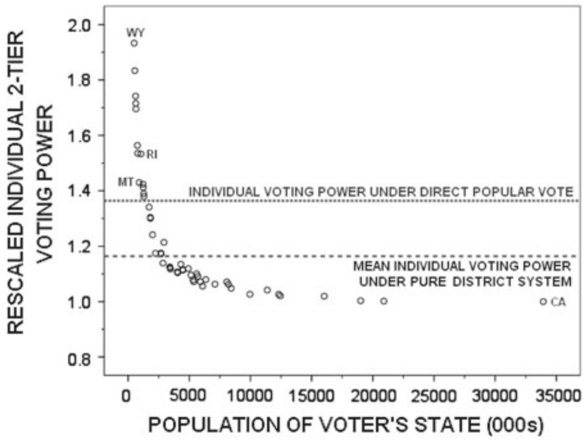


Fig. 11 Individual voting power by state population under the pure district plan

particular, the small-state apportionment advantage and the whole-number effect. The small-state advantage in apportionment carries through to voting power—for example, voters in Wyoming have almost twice the voting power as those in California, but it is substantially diluted by the Inverse Banzhaf Effect. While California districts have almost four times as many voters as the Wyoming districts, California voters have about half the voting power of those in Wyoming.

Under the Modified District Plan, two electors are elected at-large in each state and the others are elected by Congressional Districts. Individual voting power within each state is equal, because each district has an equal number of voters. All districts have equal voting power in the Electoral College, because they have equal weight, i.e., 1 electoral vote; and all states have equal voting power in the Electoral College, because they have equal weight, i.e., 2 electoral votes. But individual voting power across states is not equal, because districts in different states have different numbers of voters (due to the whole-number effect) and states with different populations have equal electoral votes.

The Modified District Plan is more complicated than it may at first appear, as the same votes are aggregated in two different ways, with the result that doubly decisive votes can be cast in three distinct contingencies: (a) a vote is decisive in the voter's district (and the district's one electoral vote is decisive in the Electoral College); (b) a vote is decisive in the voter's state (and the state's two electoral votes are decisive in the Electoral College); and (c) a vote is decisive in both the voter's district and state (and the combined three electoral votes are decisive in the Electoral College).

Moreover, because each individual vote counts in two ways, there are interdependencies in the way in which district and state electoral votes may be cast. Whichever candidate wins the two statewide electoral votes must win at least one district electoral vote as well but need not win more than one. Thus in a state with a single House seat, individual voting power under the Modified District Plan operates in just the same way as under the existing Electoral College, as its three electoral votes are always cast in a winner-take-all manner for the state popular vote winner. In a state with two House seats, the state popular vote winner is guaranteed a majority of the state's electoral votes (i.e., either 3 or 4) and a 2-2 split cannot occur. In a state with three or more House seats, electoral votes may be split in any fashion and, in a state with five or more House seats, the statewide popular vote winner may win only a minority of the state's electoral votes—that is, "election inversion" may occur at the state, as well as the national, level.

However, the preceding remarks pertain only to logical possibilities. Probabilistically, the casting of district and statewide electoral votes will be to some degree aligned in Bernoulli (and other) elections. Given that a candidate wins a given district, the probability that the candidate also wins statewide is greater than 0.5—that is to say, even though individual voters cast statistically independent votes, the fact that they are casting individual votes that count in the same way in two tiers (districts and states) induces a correlation between popular votes at the district and state levels within the same state. This correlation, which is perfect in the states with only one House seat, diminishes as a state's number of House seats increases, and therefore enhances individual voting power in small states relative to what it is under the Pure District Plan. But this correlation also makes the calculation of individual two-tier voting power far from straightforward.

The first step is to determine the probability of each of the three first-tier contingencies in which a voter may be doubly decisive. (See Miller 2013 for further details.) The probability that the district vote is tied can be calculated by the Inverse Square Root Rule, and likewise the probability that the statewide vote is tied. The conditional probability that the state vote is tied, given that the district vote is tied, is equal to the probability that the popular vote cast in all other districts in the state together is tied, which can be calculated by the Inverse Square Root Rule. By multiplying this conditional probability by the probability that the district vote is tied in the first place, we get the probability that both district and state votes are tied, i.e., the probability of contingency (c) above. The probabilities of the two other contingencies can then be determined by simple subtraction.

Having determined the probability of each contingency that makes a voter decisive in the first tier, we must calculate the probability that the single electoral vote of the district, or the two electoral votes of the state, or the combined three electoral votes of both (as the case may be) are decisive in the second tier. At first blush, it might seem that we need only evaluate the voting power of units within an Electoral College of 436 units with one electoral vote each and 51 units with two electoral votes each, but to do this ignores the interdependencies and correlations discussed earlier.[15]

[15] Banzhaf III (1968) presented calculations for the Modified District Plan that ignored these interdependencies. Had he displayed the *absolute* voting power of voters in each state, it would have been evident that mean individual voting power under the district plan (as he calculated it) exceeded that under direct popular vote, which Felsenthal and Machover (1998, pp 58–59) show is an impossibility. However, his *rescaled* voting power values are quite close to those presented here.

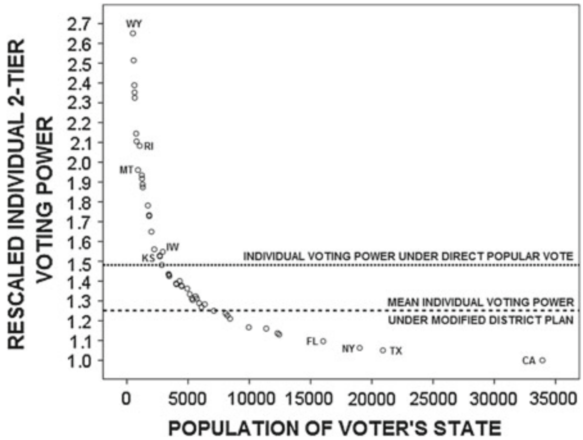


Fig. 12 Individual voting power by state population under the modified district plan

While it may be possible to proceed analytically, I have found the obstacles to be formidable and have instead proceeded by generating a sample of 1,080,000 Bernoulli elections, with electoral votes awarded to the candidates on the basis of the Modified District Plan.[16] This generated a database that can be manipulated to determine frequency distributions of electoral votes for the focal candidate under specified contingencies with respect to first-tier voting, from which relevant second-tier probabilities can be inferred.[17]

Figure 12 shows individual voting power across the states under the Modified District Plan. Voters in small states are more favored than under the Pure District Plan (because small states come closer to maintaining winner-take-all than larger states), but the Inverse Banzhaf effect within each state still attenuates the small-state apportionment advantage relative to their advantage under the Pure Proportional Plan, to which we now turn.

With sufficiently refined proportionality, the Pure Proportional Plan creates a 122-million *single-tier* (rather than a two-tier) *weighted* voting system, where the weight of individual votes is given by their state's electoral votes divided by the

[16] The simulation took place at the level of the 436 districts, not individual voters. For each Bernoulli election, the popular vote for the focal candidate was generated in each Congressional District by drawing a random number from a normal distribution with a mean of $n/2$ and a standard deviation of $0.5\sqrt{n}$ (i.e., the normal approximation of the symmetric Bernoulli distribution), where n is the number of voters in the district. (Of course, the other candidate won the residual vote.) The winner in each district was determined, the district votes in each state were added up to determine the state winner, and electoral votes are allocated accordingly.

[17] Even with the very large sample, few elections were tied at the district or state level, so the relevant electoral vote distributions were taken from a somewhat wider band of elections, namely those that fell within 0.2 standard deviations of an exact tie. (In a standard normal distribution, the ordinate at $\pm 0.2 \times$ SDs from the mean is about 0.98 times that at the mean.) It needs to be acknowledged that Fig. 12 (and Figs. 15a and 15b for the National Bonus Plan) are not as accurate as other figures, as they entail some sampling error, some other approximations (including the one just noted), and possibly other errors.

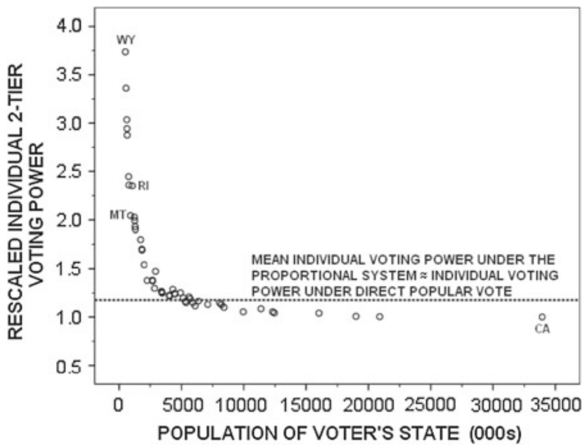


Fig. 13 Individual voting power by state population under the pure proportional plan

number of voters in the state.[18] The calculations displayed in Fig. 13 assume that proportionality is sufficiently refined to create a single-tier weighted voting system and use the Penrose Limit Theorem to justify the assumption that voting power is proportional to voting weight in this very large-*n* single-tier weighted voting system.[19] It can be seen that under this plan the small-state apportionment advantage carries through to individual voting power without the dilution evident under the Pure District Plan (or, to a lesser degree, under the Modified District Plan), because in a single-tier voting system there is no room for the Inverse Banzhaf Effect. Thus the fact that Wyoming has almost four times the electoral votes per capita as California translates without dilution into voting power for Wyoming voters that is likewise almost four times that of California voters. On the other hand, states that are relatively small but not among the smallest (with a population of about 2.5 to 5 million) are less favored relative to both the smallest states and larger states under the Pure Proportional Plan than under the Pure District Plan or (especially) the Modified District Plan. Put otherwise, the implicit "voting power by state population curve" in Fig. 13 bends more abruptly in the vicinity of the "southwest" corner of the chart than in either Fig. 11 or (especially) Fig. 12. Finally, since sufficiently refined proportionality creates what is effectively a weighted single-tier voting game (with relatively equal weights), mean

[18] The Lodge-Gossett Plan proposed in the 1950s specified that candidates would be credited with fractional electoral votes to the nearest one-thousandth of an electoral vote. As proportionality becomes less refined, this system begins to resemble the Whole-Number Proportional System. The Pure Proportional Plan has recently been reinvented as the "Weighted Vote Shares" proposal of Barnett and Kaplan (2007). Combining a precisely proportional method of *casting* of electoral votes with a precisely proportional *apportionment* of electoral votes (as discussed earlier) would give every voter equal weight and would be equivalent to direct popular vote.

[19] Banzhaf III (1968) presented similar calculations based on similar, though less explicit, assumptions.

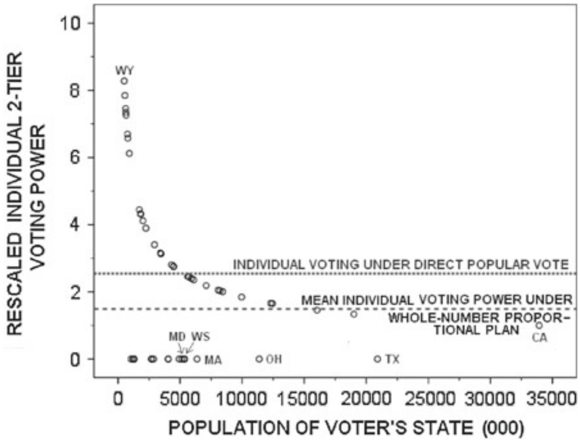


Fig. 14 Individual voting power by state population under the whole-number proportional plan

individual voting power is essentially equal to (but in principle slightly less than) individual voting power under a direct (unweighted) popular vote.[20]

The Whole-Number Proportional Plan divides a state's electoral votes between (or among) the candidates in a way that is as close to proportional to the candidates' state popular vote shares as possible, given that the apportionment must be in whole numbers. In principle, there are as many such plans as there are apportionment formulas. In addition (and as under many proportional representation electoral systems), candidates might be required to meet some vote threshold in order to win electoral votes.[21] But, in the event there are just two candidates (as we assume here), all apportionment formulas work in the same straightforward way: multiply each candidate's share of the popular vote by the state's number of electoral votes to derive his electoral vote *quota* and then round this quota to the nearest whole number in the normal manner.[22] In this two-tier system, individual *a priori* voting power is the probability that the voter casts a decisive vote within his or her state, in the sense that other votes in the state are so divided that the individual's vote determines whether a candidate gets k or $k + 1$ electoral votes from the state and that this single electoral vote is decisive in the Electoral College (where, as usual, these probabilities result from Bernoulli elections).

[20] Mean voting power under the Pure Proportional Plan (as calculated here) is 0.000072150172 versus 0.0000721502396 under direct popular vote.

[21] Colorado's Proposition 36 had no explicit vote threshold but used a distinctly *ad hoc* apportionment formula that was overtly biased in favor of the leading candidate and against minor candidates.

[22] Given three or more candidates, simple rounding does not always work, because the rounded quotas may not add up to the required number of electoral votes—hence the "apportionment problem" definitively treated by Balinski and Young (1982).

Figure 14 shows that the Whole-Number Proportional Plan produces a truly bizarre allocation of voting power among voters in different states.[23] Voters in the 17 states with an even number of electoral votes are rendered (essentially) powerless. Voters in the 33 states and the District of Columbia with an odd number of electoral votes have voting power (essentially) as if each of these states had equal voting weight (in the manner of Fig. 9). Here's why this happens.

In a Bernoulli election with fairly large number of voters, the vote essentially always is divided almost equally between the two candidates. As previously noted, the expected vote share for each candidate is 0.5 with a standard deviation of $0.5\sqrt{n}$. Consider a state with four electoral votes. For its electoral votes to be divided otherwise than 2 to 2, one candidate must receive more than 62.5 % of the vote, because 0.625×4 earns an electoral vote quota of 2.5 electoral votes, and anything below this rounds to 2. Such a state has about 500,000 voters, so the expected vote share for either candidate in a Bernoulli election is 250,000 with a standard deviation of about $0.5\sqrt{500{,}000}, = 354$ votes. Since a candidate has to receive 62,500 votes (about 175 standard deviations) above this expected vote share in order for anyone to cast a decisive vote, it is essentially guaranteed that the electoral vote will be split 2-2, giving each voter essentially zero probability of casting a decisive vote. As the even number of electoral votes increases, two things change. First, the relative vote margin required to produce anything other than an even split of electoral votes decreases. For example, in a state with 50 electoral votes, a candidate needs to get only 51 % of the vote to earn a quota over 25.5 electoral votes, anything above which rounds off to 26. At the same time, while the absolute standard deviation of the expected vote percent increases with the square root of electorate size, the relative standard deviation (expressed as a percent of the vote) decreases with the square root of electorate size. Overall, the gap between the required margin and 50 % relative to the standard deviation diminishes with electorate size, but not nearly fast enough to give voters measurable *a priori* voting power in even the largest states.

With respect to the 34 states with an odd number of electoral votes, the results are only slightly less bizarre. For (appropriate) example, consider Colorado with 9 electoral votes. Whichever candidate receives the most popular votes wins at least 5 electoral votes. But to win more than 5 electoral votes, a candidate must earn an electoral vote quota of more than 5.5 (rounding to 6), which requires a bit over 61 % of the popular vote. Even in state with 55 electoral votes (e.g., California), one candidate must win a bit over 51.8 % of the votes to win more than 28 of them. By the same considerations that applied in the even electoral vote case, the probability of achieving such margins in a Bernoulli election is essentially zero. Thus in each state with an odd number of electoral votes, effectively only one electoral vote is at stake, and the distribution of voting power is effectively the same as if electoral votes were equally apportioned among these states, thereby

[23] Similar calculations and chart were independently produced and published by Beisbart and Bovens (2008).

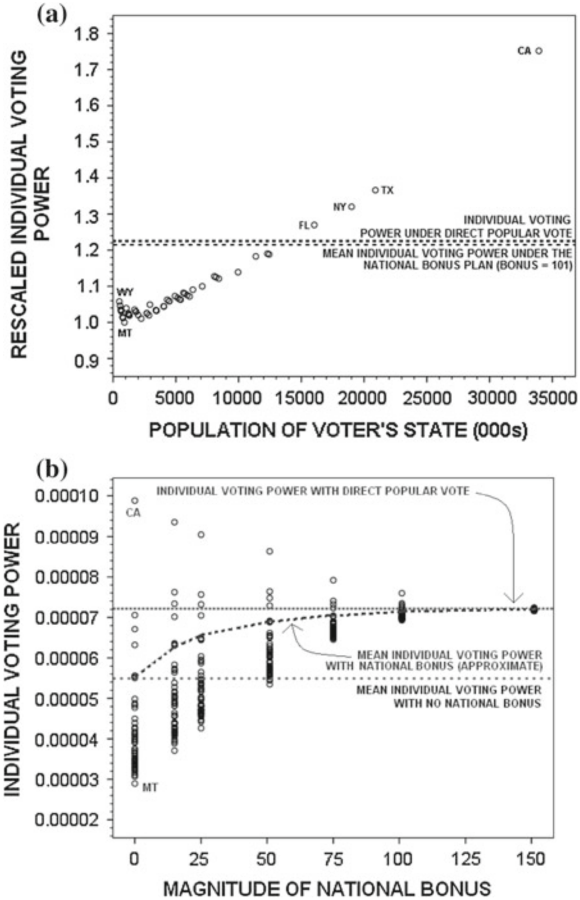


Fig. 15 **a** Individual voting power by state population under the national bonus plan (bonus of 101). **b** National bonus plan (bonus of 101)

giving a huge advantage to voters in smaller states with an odd number of electoral votes.

Finally, we take up the National Bonus Plan, focusing particularly on a bonus of 101 electoral votes. In this event, there are 639 electoral votes altogether, with 320 required for election (and ties are precluded). As with the Modified District Plan, doubly decisive votes can be cast in three distinct contingencies: (a) a vote is decisive in the voter's state (and the state's electoral votes are decisive in the Electoral College); (b) a vote is decisive in the national election (and the national bonus is decisive in the Electoral College); and (c) a vote is decisive in both the voter's state and in the national election (and the combined state and bonus electoral votes are decisive in the Electoral College).

The probabilities of the first-tier contingencies can be calculated in the same manner as those for the Modified District Plan. I then generated a sample of 256,000 Bernoulli elections, with electoral votes awarded to the candidates on the basis of the National Bonus Plan (with bonuses of varying magnitudes). Again this generated a database that can be manipulated to determine frequency distributions of electoral votes for the focal candidate under specified contingencies with respect to first-tier voting, from which relevant second-tier probabilities can be inferred.

Figure 15a displays individual voting power with a national bonus of 101 electoral votes. At first blush, Fig. 15a may look very similar to Fig. 5 for the existing Electoral College, but inspection of the vertical axis reveals that the inequalities between voters in large and small states are greatly compressed relative to the existing system. Figure 15b displays individual voting power under national bonuses of varying magnitude. A bonus of zero is equivalent to the existing Electoral College system and a bonus of 533 is logically equivalent to direct popular vote,[24] though Fig. 15b indicates that any bonus greater than about 150 is essentially equivalent to direct popular vote. Sampling error presumably accounts for the minor anomalies in this chart, but the overall patterns are clear enough. As the size of the bonus increases, voting power inequalities are compressed and mean individual voting power increases until it equals that under direct popular vote.

6 Conclusions

I conclude with a few summary points, observations, and qualifications.

Figure 16 summarizes and compares individual two-tier voting power under all Electoral College variants that entail unequal voting power. In this chart, voting power must be expressed in absolute terms, rather than be rescaled so that the

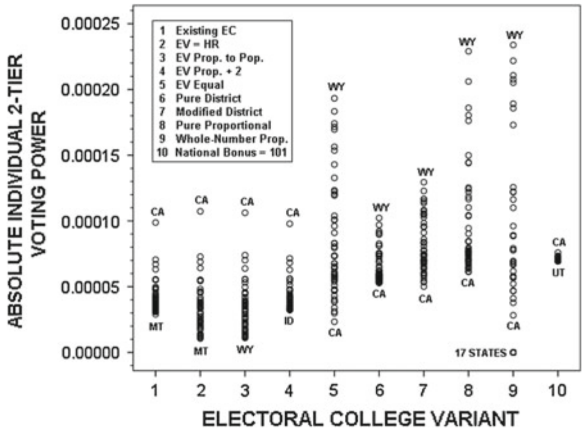


Fig. 16 Summary: individual voting power under electoral college variants

[24] Just as a statewide winner under the Modified District Plan must win at least one district, the national popular vote winner must win at least one state with at least 3 electoral votes; 533 is the smallest number B such B + 3 > 538−3.

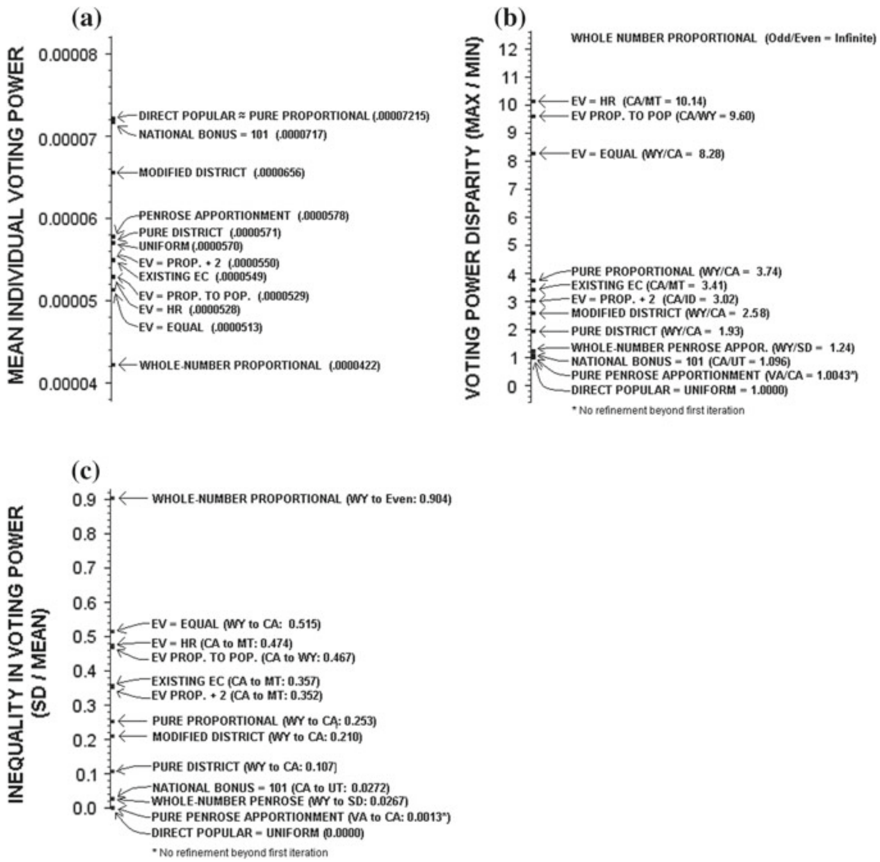


Fig. 17 **a** Mean voting power under Electoral College variants. **b** Maximum versus minimum voting power under Electoral College variants. **c** Inequality in voting power under Electoral College variants

voting power of the least favored voter is 1.00, because it makes comparisons across Electoral College variants under which different voters are least favored and the absolute voting power of the least favored voters varies. While the existing Electoral College favors voters in large states with respect to *a priori* voting power, all alternative electoral vote-casting plans would shift the balance of voting power quite dramatically in favor of voters in small states. The National Bonus Plan is a partial exception, in that it reduces the large-state advantage as the magnitude of the bonus increases and equalizes voting power given a sufficiently large bonus.

The ten columns of plotted points in Fig. 16 indicate that there are substantial differences among the plans with respect to both the mean level of individual voting power and inequality of voting power. The first point is highlighted in Fig. 17a, which ranks all variants (now including uniform and Penrose apportionments, plus direct popular election) with respect to the mean level of individual voting power that they entail. Direct popular election establishes a maximum that cannot be exceeded, but it is essentially matched by the Pure Proportional Plan and the National Bonus Plan (with a bonus of 101) does almost as well. The Modified District Plan follows some distance behind. At the lower extreme, the Whole-Number Proportional Plan, which renders a large proportion of voters powerless, ranks well below all other variants, while the remaining variants are all clustered quite closely together in the middle of the range.

Figures 17b and 17c focus on inequality of individual voting power. Figure 17b summarizes information that is also directly apparent in Fig. 16, by ranking the Electoral College variants with respect to the ratio of maximum to minimum individual voting power that they entail. This ratio is essentially infinite under the Whole-Number Proportional Plan (favoring small states with an odd number of electoral votes), and it is very high (favoring large states) when electoral votes are apportioned (whether in fractions or whole numbers) proportional to population and it is also high (but favoring small states) when states have equal electoral weights. Direct popular vote and uniform apportionment achieve perfect equality, as does pure Penrose apportionment if sufficiently refined. Whole-number Penrose apportionment does almost as well. The remaining systems are clustered fairly close together in the lower middle portion of the range. Figure 17c assesses the variants with respect to inequality of voting power more comprehensively in terms of the ratio of the standard deviation to the mean of individual voting power. The same five variants achieve perfect or close-to-perfect equality, and the Whole-Number Proportional Plan remains the extreme outlier in the other direction, though it can now be placed at a definite point on the scale. The other systems are ranked much as in Fig. 17b but are spread over a larger portion of the total range.

The analysis presented in this chapter has been static, in particular by considering Electoral College variants in turn and assuming that the manner in which states cast their electoral votes is fixed and uniform. But states are free to switch unilaterally from the existing winner-take-all system to either district plan or to the Whole-Number Proportional Plan. Therefore, it is worth observing that, in so far as states chose their mode of casting electoral votes with an eye to maximizing the power of their voters, the existing (almost) universal winner-take-all method is an "equilibrium choice"—that is, no state (or small subset of states) has an incentive to switch from winner-take-all to one of the available alternatives. For example, in the mid-1990s the Florida state legislature gave serious consideration to a proposal to use the Modified District Plan, though it ultimately rejected the proposal. The effect on individual voting power of a switch by Florida away from winner-take-all is shown in Fig. 18a (which, however, assumes a switch to the Pure District Plan, because the calculations are straightforward). Individual voting power in Florida would have been cut to about one-sixth of its previous magnitude, while the power of voters in all other states would have been slightly increased.[25] Likewise, had Colorado voters

[25] It would appear that Maine and Nebraska have been penalizing themselves in the same fashion for several decades, but the penalty for departing from winner-take-all is much less severe for

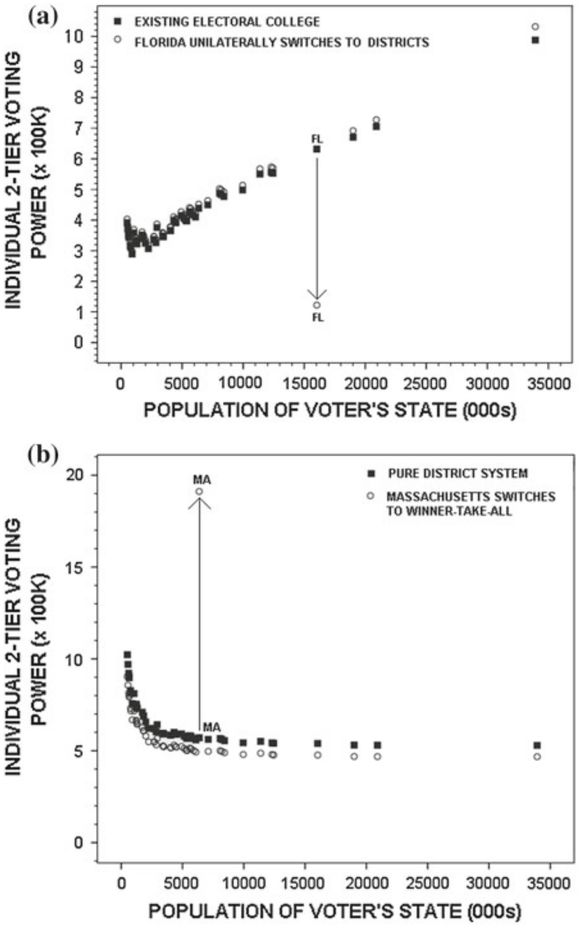


Fig. 18 **a** Individual voting power: Florida switches from winner-take-all to pure district plan. **b** Individual voting power: Massachusetts switches from pure district plan to winner-take-all

passed Proposition 36 and put the Whole-Number Proportional System into effect in their state, they would have (with respect to *a priori* voting power) in effect been throwing away four of their five electoral votes—or all of them, in the event Colorado were to gain (or lose) a House seat in the next apportionment.

Moreover, a universal winner-take-all system is not simply an equilibrium choice; it appears to be the only equilibrium, and it has strongly "attractive" as

(Footnote 25 continued)
smaller states. If Maine used the Pure District System instead of winner-take-all, the power of its voters would be cut approximately in half. Since it actually uses the Modified District Plan and is small enough that this system entails "winner-take-almost-all" (i.e., at least three of its four electoral votes), the actual reduction in voting power of Maine voters is less than this. (Another consequence of a Florida switch to districts would have been that—at least considering "mechanical" effects only—Gore would have been elected President in 2000, with no room for dispute and regardless of who won the statewide vote in Florida).

well as "retentive" properties. For example, prior to the 1800 election, Massachusetts switched from a mixed system of selecting Presidential electors to legislative appointment, which in practice meant winner-take-all for the locally dominant Federalist Party. A concerned Jefferson wrote to Monroe (cited by Pierce and Longley 1981, p 37):

> All agree that an election by districts would be best if it could be general, but while ten states choose either by their legislatures or by a general ticket [i.e., in either event, winner-take-all], it is folly or worse for the other six not to follow.

At the instigation of Jefferson and the locally dominant Republican Party, Virginia switched from the Pure District Plan to winner-take-all a general ticket for the 1800 election. If it had not done so, the Jeffersonian Republicans might easily have lost enough Virginia districts to lose the national electoral vote. Figure 18b, though using the present apportionment of electoral votes, powerfully confirms Jefferson's strategic insight in terms of individual voting power (though the voting-power rationale for winner-take-all is logically distinct from Jefferson's party-advantage rationale). Given a universal district system, Massachusetts (or any other state, but large states even more than small states) would gain substantially by switching from districts to winner-take-all. As other states follow, they also would gain but not as much as Massachusetts initially did and they would erode the initial advantage of the earlier switchers. No equilibrium is reached until all states switch to winner-take-all, even though small states would end up worse off than at the outset. Moreover, at least under the present apportionment, mean voting power would end up slightly lower than at the outset. Even if a district system were universally agreed to be socially superior (as Jefferson evidently considered it), states are caught in a kind of Prisoner's Dilemma and would not voluntarily retain (or return to) such a system, though they would happily ratify a constitutional amendment mandating it nationwide.

Finally, I should acknowledge that there are several important critiques of Banzhaf voting power measurement as applied to the Electoral College and similar two-tier voting systems (e.g., Margolis 1983; Gelman et al., 2002, 2004; Katz et al., 2004). These critiques rest fundamentally on the (indisputable) observation that Bernoulli elections are in no way representative of empirical voting patterns. But these critiques overlook the fact that the Banzhaf measure pertains to *a priori* voting power. It measures the power of states—and, in the two-tier version, of individual voters—in a way that takes account of the Electoral College voting rules *but nothing else*. As we have seen, a voter in California is about three times more likely to cast a decisive vote than one in New Hampshire in a Bernoulli election. But if we take account of recent voting patterns, poll results, and other information, a voter in New Hampshire undoubtedly has had greater *empirical* (or *a posteriori*) probability of decisiveness in recent elections, and accordingly got more attention from the candidates and party organizations than one in California. But if California and New Hampshire were both perfectly contested "battleground" states, California's *a priori* advantage would be surely reflected in its *a posteriori* voting power as well.

If it is hardly related to empirical voting power in any particular election, the question arises of whether *a priori* voting power should be of concern to political science and practice. I think the answer is yes. In particular, constitution-makers arguably should, and to some extent must, design political institutions from behind a "veil of ignorance" concerning empirical contingencies and future political trends. Accordingly they should, and to some extent must, be concerned with how the institutions they are designing allocate *a priori* voting power.

Acknowledgments I thank Dan Felsenthal, Moshe Machover, and especially Claus Beisbart for very helpful criticisms and suggestions.

References

Balinski, M. L., & Young, H. P. (1982). *Fair representation: meeting the ideal of one man, one vote*. New Haven: Yale University Press.

Banzhaf III, J. F. (1968). One man, 3.312 votes: a mathematical analysis of the Electoral College. *Villanova Law Review, 13*, 304–332.

Barnett, A., & Kaplan, E. H. (2007). A cure for the Electoral College? *Chance, 20*, 6–9.

Beisbart, C., & Bovens, L. (2008). A power measure analysis of the amendment 36 in Colorado. *Public Choice, 134*, 231–246.

Chang, P.-L., Chua, C. H. V., & Machover, M. (2006). L.S. Penrose's limit theorem: tests by simulation. *Mathematical Social Sciences, 51*, 90–106.

Felsenthal, D. S., & Machover, M. (1998). *The measurement of voting power: theory and practice, problems and paradoxes*. Cheltenham: Edward Elgar.

Gelman, A., Katz, J. N., & Tuerlinckx, F. (2002). The mathematics and statistics of voting power. *Statistical Science, 17*, 420–435.

Gelman, A., Katz, J. N., & Bafumi, J. (2004). Standard voting power indexes do not work: an empirical analysis. *British Journal of Political Science, 34*, 657–674.

Katz, J. N., German, A., & King, G. (2004). Empirically evaluating the Electoral College. In A. N. Crigler, M. R. Just, and E. J. McCaffrey (Eds.), *Rethinking the vote*. Oxford: Oxford University Press, pp. 75–88.

Kolpin, V. (2003). Voting power under uniform representation. *Economics Bulletin, 4*, 1–5.

Kuroda, T. (1994). *The origins of the Twelfth Amendment: the Electoral College in the early republic, 1787–1804*. Westport: Greenwood Press.

Lindner, I., & Machover, M. (2004). L.S. Penrose's limit theorem: proof of some special cases. *Mathematical Social Sciences, 47*, 37–49.

Margolis, H. (1983). The Banzhaf fallacy. *American Journal of Political Science, 27*, 321–326.

Martin, L (1787). "The genuine information delivered to the legislature of the state of maryland relating to the proceeding of the general convention," in M. Farand (Ed.), *The records of the Federal Convention of* 1787 (1937 rev. ed.), Vol. 3, New Haven: Yale University Press, pp. 172–232.

Miller, N. R. (2012). Election inversions by the U.S. Electoral college. In D.S. Felsenthal and M. Machover (Eds.), *Electoral systems: assumptions, and procedures*. Berlin: Springer, pp. 93–127.

Miller, N. R. (2013). "*A priori* voting power when one vote counts in two ways, with application to two variants of the U.S. Electoral College." In R. Fara, D. Leech, and M. Salles, (Eds.), *Voting power and procedures: essays in honour of Dan S. Felsenthal and Moshé Machover*. Berlin: Springer.

Neubauer, M. G., & Zeitlin, J. (2003). Outcomes of presidential elections and House size. *PS: Politics and Political Science, 36*, 221–224.

Peirce, N. R., & Longley, L. D. (1981). *The people's president: the Electoral College in American history and the direct vote alternative*. New Haven: Yale University Press. revised.

Penrose, L. S. (1946). The elementary statistics of majority voting. *Journal of the Royal Statistical Society, 109*, 53–57.

Riker, W. H. (1986). The first power index. *Social Choice and Welfare, 3*, 293–295.

Do Voting Power Considerations Explain the Formation of Political Coalitions? A Re-Evaluation

Vincent C. H. Chua and Dan S. Felsenthal

1 Introduction

In a multi-party system in which no single party controls a decisive majority in the legislature, the party charged with forming a new government sometimes has the option of governing as a minority government. Such a course of action will preserve the party's control over all ministerial portfolios and its legislative agenda but it will also mean an arduous uphill battle to see the agenda through since, as a governing party, it does not control the required majority in the legislature. Therefore, in many instances, forming a majority coalition government[1] is a preferred arrangement. Here, it will usually have the liberty to decide which other party or parties may join the governing coalition.

Involving other parties in a coalition government, however, entails a cost and while the leading party still has control over the distribution of ministerial portfolios and plays the key role in shaping the legislative agenda, it will have to cede some measure of control to other coalition members. The vulnerability of the governing coalition to defection will also feature in the political calculations. Citing these considerations in a 1995 interview, Aumann (see van Damme 1998)

[1] In this chapter, the term *coalition* is regarded as synonymous with the term *alliance* in the sense of Felsenthal and Machover (2002, 2008) and should be understood as such.

V. C. H. Chua (✉)
SIM University, 461 Clementi Road, Singapore 599491,
e-mail: vincentchua@unisim.edu.sg

D. S. Felsenthal
University of Haifa, Haifa, Israel
e-mail: dfelsenthal@poli.haifa.ac.il

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_23, © Springer-Verlag Berlin Heidelberg 2013

hypothesized that, when selecting its coalition partners, a coalition leader will act in a manner that will maximize its Shapley (1953) value.[2]

In Chua and Felsenthal (2008), we subjected this hypothesis and three variants to empirical testing using two sets of data: historical election data tabulated in de Swaan (1973) for eight European countries and for Israel; and detailed election data covering all Israeli elections from 1949 through 2006. In this connection, consistent with Aumann's elaboration of his hypothesis, we defined a winning coalition as any coalition that controls a simple majority of seats in the legislature and which included the party charged with forming the coalition. But, as our investigation revealed, neither the hypothesis nor its variants managed a level of predictive success that is anywhere near that achieved by the closed minimal range theory[3] or, for that matter, by the minimum size principle,[4] leading us to summarily reject the hypothesis and its variants.

The behavioral considerations that led to Aumann's hypothesis were, however, intuitive and compelling and have long been the subject of inquiry by political scientists who made a clear distinction between those considerations that are 'office-seeking' and those that are 'policy-seeking'.[5] In advocating the Shapley value as the relevant summary measure for the different behavioral considerations, Aumann had implicitly emphasized the office-seeking element, the element that is associated with the notion of power as *prize* or *P*-power.[6] As explained in Felsenthal and Machover (1998, p. 18), the use of such a measure implicitly regards the coalition formation process as an *n*-person bargaining game involving a fixed purse—the *prize of power*. Political parties with representation in the legislature are, from this perspective, viewed as participating in this bargaining game with the winning coalition capturing the entire pie which it then proceeds to divide among its members.

But, in an actual legislature, it is unclear if this office-seeking motive indeed occupies center-stage position in the 'mind' of the party charged with forming the government, and even when it does feature in an important way, whether it is not merely because control over key political portfolios is instrumental to the smooth implementation of the party's policy agenda. After all, when a party is voted into power, it is usually based on the merits of the party platform and on the premise

---

[2] In the context of simple voting games, the Shapley value is usually referred to as the Shapley-Shubik index. See Shapley and Shubik (1954).

[3] Briefly, on a policy scale in which any two adjacent parties are regarded as one unit distance apart, the range of a coalition is the distance between the two parties in the coalition whose positions on the policy scale are furthest apart. The minimal range theory, due to Leiserson (1966) and to Axelrod (1970), hypothesizes that the coalition formed will be of minimal range.

[4] This hypothesis is due to Riker (1959, 1962) and to Gamson (1961). Gamson uses the term *cheapest winning coalition* to describe the winning coalition controlling the smallest total number of seats (or votes) while Riker uses the term *coalition of minimal size*.

[5] For a discussion of these different motivations, the reader is referred to Laver and Schofield (1990), especially Chap. 3.

[6] For a thorough discussion of *P*-power or power as *prize*, see Chap. 6, Felsenthal and Machover (1998).

that there is a good chance the party will deliver on its promises, which in turn puts pressure on the party to perform. From a realist perspective, therefore, moving the party agenda forward and/or influencing the legislative outcome in this direction would seem to be a more natural and central pre-occupation of any party leading the coalition formation process. Additionally, this desired influence ought to be over the period of its tenure in office, highlighting a certain permanency in the arrangement that is being sought.[7] To differentiate this type of coalitional arrangement from those that are one-off or of an *ad hoc* nature, Felsenthal and Machover (2008) have aptly termed such arrangements *alliances*.

In an alliance, the leading party is viewed as participating in a composite voting game, one involving members within the governing coalition and the other involving the entire legislature and the influence that the leading party has in the legislature is the product of its influence within the governing coalition and the influence that the governing coalition, as a bloc, has in the legislature. Since, in this chapter, we are only concerned with the formation of majority coalitions under the simple majority rule, once formed, the governing coalition will have absolute influence over the legislative outcomes. Thus the influence the leading party has within the legislature is the same as its ability to influence the way the bloc will cast its vote on each issue in the legislature. If the 'internal' decision rule for the bloc is the simple majority rule, having a majority within the bloc will confer absolute influence upon the leading party.[8]

With the preceding in mind and as we proceed with our continuing investigation into the role of voting power considerations in the political coalition formation process, it appears reasonable for us to postulate an alternative hypothesis: that the party charged with forming the government will seek to enter into an arrangement that will maximize its absolute influence or (I-) power in the legislature. And, the measure of I-power that is most relevant to our purpose is naturally the Penrose measure of absolute voting power, not the *game-theoretic* Shapley value as suggested by Aumann.

At this juncture, it is important to emphasize that, from our vantage point and consistent with casual observations, when the legislative vote goes in the direction of the ruling majority coalition, it does not preclude the possibility that the legislative outcome could also be favorable and beneficial to parties not within the ruling coalition. This clearly differentiates our vision of the coalition formation process from the *game-theoretic* view implicit in Aumann's approach in which those that are not members of the ruling coalition will end up with a zero expected payoff.

Using our approach, the empirical findings are considerably more positive and suggest to us that the Penrose measure may play a useful role in furthering our

---

[7] See, for instance, Coleman (1971) for a statement to this effect.

[8] This last point is another important departure from Aumann's approach. In Aumann's approach, the Shapley-Shubik value of the leading party in the legislature is computed using a quota that is set equal to a simple majority of the entire legislature as opposed to a simple majority of the size of the alliance.

understanding of the coalition formation process.[9] We report the details of our investigation below beginning with an outline of the Penrose measure and an example of how we have proceeded in computing the absolute influence of the leading party in the legislature. A description of the main hypothesis and its three variants follows. We then provide a brief description of the data sets employed and a discussion of our empirical findings. Some closing comments conclude the chapter.

2 The Penrose Measure of Absolute Voting Power or Influence

As noted in Felsenthal and Machover (2004), the intuition underlying Penrose's measure may be summarized by two simple yet fundamental ideas. The first, simply put, says "the more powerful a voter is, the more often will the outcome go the way s/he votes". Thus it is natural that the (a priori) power of a voter v should be directly related to the proportion of all possible divisions of the legislature in which s/he is successful. Let this proportion be denoted r_v. The second emphasizes that even a dummy will, through sheer luck, find itself on the winning side in half of all divisions. Isolating the pure luck element from the true influence yields a measure of the voter's influence that takes the following form as proposed in Penrose (1952): $\psi_v = 2r_v - 1$.

$\psi_v[W]$ is the Penrose (*I*-) power of voter v in some given simple voting game W, and as pointed out in Felsenthal and Machover (2004), it also indicates the probability that voter v is decisive. It is easy to verify that under the Penrose measure, a dummy will be assigned a value of zero whereas a dictator will be assigned a value equal to unity. Furthermore, if party v is a member of an alliance S, then its overall Penrose *I*-power in the legislature is given by its Penrose measure within the alliance, reflecting its a priori ability to influence how its fellow alliance members vote in the legislature, multiplied by the alliance's Penrose *I*-power in the entire assembly.[10] In simple terms, the coalition leader derives power or influence through a two-stage process: first, given the alliance's internal decision rule Ws, the greater the coalition leader's ability to influence the collective position of the alliance, the greater is its *internal* influence. Second, the more influential the alliance is in the

[9] At this juncture, one could argue that although the Penrose measure and the Shapley-Shubik index are not co-monotonic, departures are rare and therefore the Shapley-Shubik index will do just as well under the alternative definition of the quota used in the present investigation. This is indeed the case as our own earlier investigation has revealed. But it misses the point of our approach which emphasizes power as influence together with an associated calculus as opposed to power as prize in the game-theoretic sense with a different associated calculus.

[10] As Felsenthal and Machover (2002, 2008) have explained, denoting the power of party v within the alliance S by $\psi_v(W_s)$ and the power of the alliance in the assembly as $\psi_{\&s}(W|\&_s)$, then the overall power of party v which is a member of the alliance and is denoted $\psi_v(W\|W_s)$ is equal to $\psi_v(W_s) \cdot \psi_{\&s}(W|\&_s)$.

legislature, the greater will be the *external* influence the coalition leader will have in the legislature through the joint action of the alliance members.

In using the Penrose measure in our investigation, we have made specific assertions relating to the decision process within the legislature and within the majority coalition. Specifically, to secure the passage of a bill within the legislature, a simple majority of the total number of votes in the legislature suffices. As noted in our introductory remarks, our concern in this investigation is with the formation of majority coalitions. It is thus clear that no party outside the majority coalition wields any influence as long as the coalition remains a cohesive bloc that controls a majority of seats in the legislature. Only parties in the majority coalition wield power or influence and, as a bloc, the bloc wields absolute influence. In our investigation, we have also chosen the simple majority rule with a quota defined by the size of the majority coalition as the internal decision rule. We illustrate the computation of the measure below with an example before proceeding to a description of our hypotheses.

Example 1 (*The* 1920 *Danish Assembly*): Consider the following 4-party situation denoted by [71; 42, 17, **49**, 28]. In this example, the required majority to secure the passage of a bill is 71.[11] The four parties in the legislature, ordered according to their position on the left–right ideological continuum, are respectively the Social Democrats with 42 votes, the Danish Social Liberal Party with 17 votes, the Liberals with 49 votes, and the Conservative People's Party with 28 votes—and the party asked to form the government is the Liberals, the party with the largest number of votes. An alliance with the Social Democrats will give the Liberal-led coalition a combined weight of 91 and hence a majority in the legislature. If this alliance were to form, then under the procedure that we have adopted, the Penrose measure for the leader, the Liberals, would be 1. In contrast, if a majority alliance comprising the Social Democrats, the Danish Social Liberal Party and the Liberals were to form with a combined weight of 108 votes, the Penrose measure of the leader would be 1/2 since, with the inclusion of the Danish Social Liberal Party, the Liberal party would lose its decisive absolute majority within the majority coalition.[12]

[11] The data are taken from de Swaan (1973, p. 269). A simple majority of 71 implies that there were at that time 140 members in the Danish parliament. However, the total number of seats of the four listed parties is only 136. This gap is explained by the fact that de Swaan ignored small parties (or independents) who controlled no more than 2.5 % of the seats in parliament because such parties were very seldom included in governmental coalitions.

[12] The Penrose measure of a priori voting power of a voter v, denoted ψ_v, is equal to the number of winning coalitions in which v is critical divided by the number of coalitions to which v belongs. A coalition is winning if it has sufficient votes to pass a decision, otherwise it is losing. A voter is critical if his defection from a winning coalition renders it losing, or if his joining a losing coalition renders it winning. For n voters there are altogether 2^n coalitions (or bi-partitions) of which every voter belongs to 2^{n-1} coalitions. In the above example let us consider the alliance consisting of three members whose weights are 42, 17, and 49. This is a winning alliance (coalition) because its combined votes (weights) is 108—which exceeds the quota of 71. If we assume that the internal decision rule of this alliance is a simple majority of its members' weights (55), then there are altogether four (internal) coalitions within this alliance in which the member with weight 49 (the coalition leader) exists—{49}, {49, 17}, {49, 42}, {49, 42, 17}—of which this member is critical in the second and third coalitions. Consequently the coalition leader's

3 The Main Hypothesis and Variations

3.1 The Main Hypothesis

When charged with forming a government, we hypothesize that the party so appointed will act in a manner that will maximize its chance of implementing its policy agenda and accordingly will act in a manner consistent with the maximization of its Penrose (*I*-) power, a measure that is now widely regarded as a reasonable index for measuring the a priori absolute influence of a party in an *n-party* decision-making situation with a given quota. For convenience, we shall refer to this hypothesis as the *Maximal PI hypothesis*. Applying this hypothesis to the situation described in Example 1, it is clear that the Liberals, the party charged with forming the government, will prefer the first arrangement in which its measured a priori Penrose power is unity to the second in which the corresponding Penrose value is only 1/2 and if there are no other arrangements that will confer 'dictatorial' power on the Liberals, then the hypothesis would predict that the coalition comprising the Social Democrats and the Liberals will form. This, however, is not the case. There are two other arrangements that will also confer 'dictatorial' power on the coalition leader, specifically, the coalition comprising: (a) the Danish Social Liberal Party, the Liberals and the Conservative People's Party; and (b) the Liberals and the Conservative People's Party. This observation immediately brings to the fore a basic difficulty with our hypothesis in its present form: it lacks parsimony, a weakness that it shares with Aumann's hypothesis. A second example drawn from the 1936 Swedish election data reinforces this point.

Example 2 (*The* 1936 *Swedish Assembly*): Consider the 5-player weighted voting game given by [116; 6, **112**, 36, 27, 44]. In this game any coalition which controls at least 116 votes is winning and the parties, ordered according to their position on the left–right ideological continuum, are respectively the Communist Party of Sweden (6 votes), the Social Democrats (112 votes), the Farmers' League (36 votes), the Liberal Party (27 votes) and the Right-wing Conservative Party (44 votes).[13] In this instance, the party charged with forming the government is the Social Democrats which controlled 112 votes in the legislature. There are altogether 15 majority coalitions that the Social Democrats could consider when forming a majority government. In all except the grand coalition, however, the coalition leader has an absolute majority and thus in 14 out of the 15 instances, its Penrose *I*-power is unity.

(Footnote 12 continued)
internal (direct) voting power in this case is $2/4 = 0.5$. But since the alliance as a whole controls an absolute majority of the votes within the parliament—and hence is a dictator whose a priori voting power is 1—it follows that the overall (indirect) a priori voting power of the coalition leader in this case is equal to its internal voting power within the alliance multiplied by the alliance's power within the parliament, i.e., $0.5 \cdot 1 = 0.5$.

[13] The data are taken from de Swaan (1973, p. 260) who ignored in this case small parties controlling together five seats.

The predicted set in this instance thus comprises 14 out of the 15 possible majority coalitions, indicating clearly a lack of parsimony in the theory and rendering the prediction from the theory of little value.

3.2 Three Variations

From the perspective of developing a predictive theory of coalition formation, it is important that one has a theory that provides a reasonably precise prediction as to the likely outcome. Parsimony—in terms of the size of the predicted set of likely coalitions—is an important attribute of a good predictive theory and, therefore, to avoid the profusion of possibilities under the maximal *PI* hypothesis, we considered three variations of the main hypothesis while maximizing, minimizing, or holding constant one or more additional variables.

3.2.1 Variation I: Restriction to Closed Coalitions

In this variation, we consider a restriction of the domain to the set of *closed* or *ideologically connected* coalitions. By a 'closed' coalition, we mean that there are no ideological "gaps": if two parties belong to the coalition then any party that lies ideologically between them belongs also to the coalition; otherwise the coalition is 'open'. Returning to Example 1, of the three majority coalitions involving the coalition leader and in which the leader has the maximal Penrose value of 1, only two are closed.

The motivation for introducing this restriction is intuitive. In a multi-party system, parties that are diametrically opposed in ideology do not usually enter into a political alliance except when the situation demands it, such as in matters that involve the national security. Under ordinary circumstances, an alliance of this nature is likely to increase the cost of the arrangement to the leading party and without doubt significantly increase the vulnerability of the alliance to defection. Thus it is unlikely that a party that is charged with forming a governing coalition under normal conditions will enter into such an arrangement even though such a union may result in the highest value for the leader's Penrose *I*-power. Corroborative evidence indicating a preference by political parties for closed coalitions is provided in de Swaan (1973, p. 148); out of 108 majority coalitions in the nine countries which he investigated, 85 (or 79 %) of these were closed coalitions.

At the operational level, the restriction to closed coalitions will help to increase the parsimony of the hypothesis which, under the given restriction, we shall refer to as the *Closed Maximal PI hypothesis*.[14]

[14] It is important to make the distinction between: (a) restricting the domain to closed coalitions and picking an arrangement that is maximal from the restricted set; and (b) picking an arrangement from the maximal set that satisfies the restriction. Although (a) and (b) may pick the same arrangement, it is the former that is employed in our analysis in order to test the hypothesis

3.2.2 Variation II: Closed Maximal PI Coalitions of Minimum Size

Even with the domain restricted to the set of closed coalitions, the size of the predicted set may still be large. We thus consider a further restriction of the domain to only those closed and maximal arrangements that are also of minimum size. This will yield predicted sets that are subsets of those obtained under Variation I and, like the Gamson-Riker *minimum size principle*,[15] will yield predicted sets that are often singletons or that involve a relatively small number of coalitions. We shall refer to this variation of the hypothesis as the *Closed Maximal PI Minimal Size hypothesis*.[16]

3.2.3 Variation III: Closed Maximal PI Coalitions of Minimal Range

As an alternative to the restriction of the predicted set to only those closed maximal coalitions that are of minimal size, we also consider the restriction of the predicted set to the subset of closed maximal coalitions that are of *minimal range*. As noted previously, on a policy scale in which any two adjacent parties are regarded as one unit of distance apart, the range of a coalition is the distance between the two parties in the coalition whose positions on the policy scale are furthest apart. Inclusion of this variation is largely motivated by de Swaan's finding that Leiserson's (1966) and Axelrod's (1970) closed minimal range hypothesis appears to fit the historical data rather well. Of the 85 closed coalitions that were observed in de Swaan's study, 55 were of minimal range (see de Swaan 1973, p. 148). It also adds a dimension to the optimization process emphasizing, in addition to power, the desirability for homogeneity in the ideological positions of coalition members, an idea that is also rather intuitive. We shall refer to this other variant as the *Closed Maximal PI Minimal Range* hypothesis.

Summarizing, in addition to considering the maximal *PI* hypothesis, we shall also consider in our investigation: (a) the variation that restricts the maximization process to only those majority coalitions that are ideologically closed; (b) the variation that applies a further restriction to the predicted set to include only those with minimal size; and (c) the variation that further restricts the prediction of the Closed Maximal *PI* version to the subset that is of minimal range. While (b) may

(Footnote 14 continued)
that the maximization of the leader's voting power in forming a governmental coalition is an important consideration only (or mainly) if the coalition is closed.

15 A minimum size coalition is a winning coalition that controls no more seats in the legislature than any other winning coalition. Both Gamson (1961) and Riker (1962) predicted that the (winning) coalition that will actually form in coalition games is likely to be of minimum size.

16 Such an approach is however, distinct from the *minimum size* principle and it also overcomes the objection raised by Aumann concerning coalitional stability when the *minimum size* principle is invoked.

be regarded as an analogue to the Gamson-Riker minimum size principle, (c) may be viewed as the analogue to the Leiserson-Axelrod closed minimal range theory.

4 Data, Analysis and Findings

4.1 Data Sources

In testing our main hypothesis and its different variants, we made extensive use of the tabulation of election outcomes provided in de Swaan's (1973) book *Coalition Theories and Cabinet Formations*. This data tabulation covers nine countries and selected historical parliamentary elections: Denmark (1918–1971), Sweden (1917–1970), Norway (1933–1936; 1965–1969), Israel (1949–1969), Italy (1946–1972), The Netherlands (1918–1972), Finland (1919–1972), Germany's Weimar Republic (1919–1932), and France's Fourth Republic (1945–1957). For our purpose, de Swaan's data is particularly useful because, for each election, in addition to listing the number of seats controlled by parties which gained at least 2.5 % of the seats in an assembly, the parties have also been ordered along the left–right ideological continuum. These rank-orderings of parties render the task of identifying closed majority coalitions in each election considerably less onerous.

Differing somewhat from de Swaan's study, however, our concern is with the formation of original coalitions, that is, coalitions that are formed immediately following a general election. Interim coalitions are not included in our analysis because their formation may be due to reasons which clearly cannot be attributed to the leading party's desire to maximize its a priori voting power, for example, the death of the former prime minister, the disintegration of one of the coalition's parties, or the defection of some members from one party within the coalition to another. In implementing our empirical analysis, we also considered only those winning coalitions that included the coalition leader which we identify *ex post* as the party of the prime minister. These considerations have meant that only 65 original coalitions across the nine countries are included in our analysis.

In part, because de Swaan (1973, p. 237) has identified Israel as "a difficult country for the theories" and, in part, because, in the case of Israel, very small parties were sometimes included in governmental coalitions, we have performed additional analysis covering all 18 Israeli elections from 1949 to 2006. In this analysis, we included all parties that gained representation in Israel's parliament. This is in contrast to de Swaan's analysis of Israel that included only elections up to 1969 and covering only those parties that controlled more than 2.5 % of the seats in the various parliaments.

4.2 Evaluating the Worth of a Theory

No theory is expected to correctly predict the outcome all of the time. In some contexts, a theory that performs marginally better than chance may be considered a reasonably good theory if other competing theories can do no better. Intuitively, when a restriction is placed on the set of admissible coalitions resulting in predicted sets that are more precise, it would appear that the frequency of obtaining a correct prediction is likely to be lower compared with the case when the domain is unrestricted. This is certainly the case when the restriction results in predicted sets that are proper subsets of the unrestricted predicted sets. This does not however, automatically render the restricted theory a poorer predictive theory. Conversely, for the same election, different theories may give rise to predicted sets that differ considerably in terms of size, and the theory that gives rise to a larger predicted set will naturally have a better chance of correctly predicting the outcome. But such a theory is not necessarily a better predictive theory. Somehow, the tradeoff between the probability that the actual outcome is included in the predicted set and the parsimony of the predicted set will have to enter the calculus in determining which theory should be preferred.

In determining the worth of each theory in our analysis, we have kept in mind this tradeoff. Instead of attempting a direct comparison of the competing theories, our approach is to compare the different theories with their respective randomized counterparts so that the evaluation of each theory is carried out on a level playing field. For the de Swaan data set, given a reasonably large sample size of 65, we carry out this evaluation by invoking the Central Limit Theorem. In the case of our detailed investigation of Israel in which only 18 elections are involved, such a procedure would be inappropriate. Here, we compute the exact probability mass function under each theory to facilitate our assessment of the theories. These two evaluation procedures are outlined below.

Given the distribution of seats controlled by the various political parties in the i-th election, ($i = 1,2, \ldots, n$), let N_i denote the number of possible winning coalitions[17] that include the leading party, that is, the party charged with forming the government. If the leading party is successful in forming the government, the outcome will necessarily be one of these winning coalitions. Let S_{ij} denote the number of winning coalitions in the predicted set under the j-th theory ($j = 1,2, \ldots, m$). If $S_{ij} = 1$, then the theory in question effectively makes a unique prediction as to the government that will form in the i-th election. If the actual outcome coincides with the predicted outcome, we consider the theory as "successful" or having produced a consistent prediction; otherwise we consider the theory as having "failed".[18] If the prediction of the theory in question is no better than that of a pure chance mechanism, then the

[17] In this chapter, a winning coalition is defined as one that controls over half of the seats in the legislature.

[18] In a more general context, a prediction is regarded as consistent as long as the actual outcome is included in the predicted set.

probability of a consistent prediction would be $1/N_i$, implying that each of the N_i winning coalitions is equally likely. More generally, if $N_i > S_{ij} > 1$, this probability, which we now denote by P_{ij}, would equal S_{ij}/N_i. Needless to say, when $S_{ij} = N_i$, the predicted set coincides with the set of all winning coalitions that include the leading party and $P_{ij} = 1$. In this instance, the actual outcome is definitely contained in the predicted set. But, a theory that does this would neither be interesting nor useful.

As the distribution of seats among political parties differ across the n elections, the size of the set of winning coalitions for each of these n elections (N_1,N_2, …, N_n) and the corresponding sizes of the predicted set under the j-th theory (S_{1j},S_{2j}, …, S_{nj}) cannot be expected to remain constant. Consequently, if the prediction of the j-th theory were no better than a chance mechanism, the probability (P_{ij}) that the prediction of the j-th theory is consistent with the actual outcome will also vary across the n elections. As long as the outcomes are independent across elections, the expected number of consistent predictions under the j-th theory will be given by $\sum_i P_{ij}$ with associated standard error $[\sum_i P_{ij}(1\text{-}P_{ij})]^{1/2}$. We refer to the quantity $\sum_i P_{ij}$ as the randomized mean for the j-th theory.

For *sufficiently large n,* the Central Limit Theorem for independent random variables postulates that the expected number of consistent predictions will be *approximately normally distributed* with mean $\sum_i P_{ij}$ and standard error $[\sum_i P_{ij}(1 - P_{ij})]^{1/2}$.[19] In our analysis of de Swaan's data, we exploit this result when evaluating the worth of the j-th theory by comparing the actual number of consistent predictions obtained under the j-th theory with that obtained under its randomized counterpart. More precisely, we measure the standardized deviation of the actual number of consistent predictions of the j-th theory from its randomized mean.

Our supplemental analysis of the Israeli elections, however, requires a slightly different approach. First, the data set comprises only 18 elections, a sample size that is too small to invoke the Central Limit Theorem. Additionally, the probability of a consistent prediction (P_{ij}) in each of these 18 elections is close to zero making the actual probability mass function severely skewed. While there are a number of alternative approaches that one may consider when evaluating the worth of each theory, we opted in this chapter to compute the exact probability mass function for the randomized scheme under each of the theories, noting that the probability mass at k, the number of consistent predictions under the *j-th* theory, is given by:

$$P_j(x = k) = \sum_{a_k \in A_k} \left(\prod_{i \in a_{k1}} P_{ij} \prod_{i \notin a_{k0}} (1 - P_{ij}) \right)$$

[19] The Central Limit Theorem for independent variables states the following. Let $X_1, X_2, \cdots$ be a sequence of independent random variables having respective means and variances $\mu_i = E(X_i)$, $\sigma_i^2 = Var(X_i)$. If (1) the X_i are uniformly bounded; that is, for some M, $P\{|X_i| < \mathrm{M}\} = 1$ for all i, and (2) $\sum_{i=1}^{\infty} \sigma_i^2 = \infty$, then $P\left\{ \frac{\sum_{i=1}^{n} (X_i - \mu_i)}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}} \leq a \right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} \mathrm{e}^{-x^2/2} dx \; as \; n \to \infty$. See Ross (2002).

In this expression, A_k denotes the set of situations that gives rise to k consistent predictions and a_k is an element of this set. The cardinality of the set A_k is ${}^{n}C_k$ (i.e., the number of combinations of size k one can extract out of the set of n elections) and the summation in the expression is over each of these ${}^{n}C_k$ different situations. In each of these situations, the set a_{k1} which has cardinality k refers to the set of elections in which the theory has produced a consistent prediction. The set a_{k0} with cardinality $(n - k)$, on the other hand, refers to those elections in which the actual outcome is not in the predicted set.

As an illustration, consider a sequence of three elections, that is, $n = 3$. For the *j*-th theory, suppose P_{1j}, P_{2j} and P_{3j} are respectively 1/2, 1/3 and 1/4. If this theory is no better than a chance mechanism, there will be ${}^{3}C_0$ or exactly one instance in which it will fail to produce a consistent prediction. This is the case when $k = 0$ and denoting a consistent prediction by the letter *s* and an inconsistent prediction by the letter *f*, the sequence of outcomes referred to in this instance is *fff*. The probability that this occurs is given by $P_j\ (x = 0) = (1 - P_{1j})(1 - P_{2j})(1 - P_{3j}) = 1/2 \cdot 2/3 \cdot 3/4 = 1/4$. Similarly, there will be ${}^{3}C_1$ or exactly three instances in which the theory will produce exactly 1 consistent prediction. These three instances are associated with the outcome sequences *sff*, *fsf*, and *ffs*. Thus the probability that the theory achieves exactly one consistent prediction, $P_j\,(x = 1) = P_{1j}\,(1 - P_{2j})(1 - P_{3j}) + (1 - P_{1j})\,P_{2j}(1 - P_{3j}) + (1 - P_{1j})(1 - P_{2j})P_{3j}$ and this equals $1/4 + 1/8 + 1/12 = 11/24$. It can be similarly verified that $P_j\,(x = 2) = 1/4$ and $P_j\,(x = 3) = 1/24$. As always, these probabilities sum to unity as the different scenarios are both mutually exclusive and exhaustive.

With the exact probability mass function for each theory in hand and supposing the number of consistent predictions under the *j-th* theory is k_j^*, we are able to calculate the exact probability that this or a larger number of consistent predictions will be observed under its randomized counterpart scheme, i.e., we can calculate $P_j\left(x \geq k_j^*\right)$. This, in turn, will enable us to test the null hypothesis that the theory in question is no better than a pure chance mechanism. A large value for $P_j\left(x \geq k_j^*\right)$ is indicative that this is indeed the case, whereas a very small value for $P_j\left(x \geq k_j^*\right)$ provides strong evidence against the null hypothesis, indicating that the theory in question strongly outperforms its randomized counterpart in its predictions and thus is a candidate deserving further consideration.

4.3 Empirical Findings

Detailed computational results of our empirical work based on the de Swaan data set are appended as Annex Tables 2 and 3. In Annex Tables 4 and 5, we list the details from our analysis of the separate set of 18 Israeli elections and the exact probability mass functions that we employed. In all these tables there are entries of the type *x*/*y* in the columns of the various tested theories. Thus, for example, in the first row of

Annex Table 2 under the maximal *PI* column appears the entry 16/98. This entry should be interpreted thus: of the total 98 possible coalitions in which the leader's Penrose measure is maximized over all 17 elections considered, 16 coalitions were actually formed. The same interpretation applies, *mutatis mutandis*, to all other entries of this form. The key findings from our analyses are summarized in Table 1.

The *Maximal PI* hypothesis performs exceedingly well securing the largest number of predictive successes. The number of consistent predictions is highest of all theories considered: 43 consistent predictions out of 65 elections for de Swaan's data set (cf. Annex Table 2) and 11 consistent predictions out of 18 elections for the Israeli 1949–2006 data set (cf. Annex Table 4). But as is readily verified by examining the columns labeled *Maximal PI* in Annex Tables 3 and 4, in many elections the size of the predicted set under this hypothesis is large relative to the number of winning coalitions. Take for instance the 1936 election to the Swedish Assembly: of a total of 15 winning coalitions, 14 (or 93 %) are included in the predicted set under this hypothesis. Likewise, for the 1969 Israeli election reported in Annex Table 4: of 4,069 possible winning coalitions, 3,905 (or 96 %) are contained in the predicted set. Not surprisingly, therefore, despite the large number of consistent predictions, its performance for the de Swaan data set is only 4.1831 standard errors above its randomized mean, the lowest obtained of all the hypotheses considered. Similarly, in the case of the more detailed Israeli elections data, despite having achieved 11 out of 18 predictive successes, its performance is among the worst of all hypotheses that managed to score at least one success.

Inferring from the exact probability mass function for its counterpart randomized scheme as detailed in Annex Table 5, the probability of observing 11 or more consistent predictions under the randomized scheme is 0.051, indicating a 5.1 % chance that this event will occur. Thus, while at the 90 % confidence level, the conclusion that its performance is significantly different from its purely randomized counterpart cannot be rejected, at the 99 % confidence level, this position will be squarely rejected. In sum, under the given circumstances, the lack of parsimony of the maximal *PI* hypothesis makes it an unattractive candidate as a predictive theory of coalition formation.

For the purpose of comparison, we also investigated the predictive performance of the *cheapest coalition* or *minimum size* principle. Like the maximal *PI* hypothesis, it is simple and the informational requirements are similar. But unlike the maximal *PI* hypothesis, it is extremely parsimonious and often produced predicted sets that are very small relative to the number of winning coalitions. This principle appears to work reasonably well for the de Swaan data set, managing 26 consistent predictions in 65 elections or 10.484 standard errors above its randomized mean. For the detailed analysis of the 18 Israeli elections, however, it predicted the outcome of only one election correctly. Additionally, when its performance is compared with those achieved under other variations of the maximal *PI* hypothesis, which we shall discuss shortly, this theory appears to be a dominated theory.

Unlike the hypotheses just discussed, all three variants of the maximal *PI* hypothesis considered in this investigation require that the selected coalitions be closed on the left–right continuum. This imposes the additional informational

Table 1 Summary test statistics on the predictive accuracy of the theories

	Closed	Maximal *PI*	Closed maximal *PI*	Closed maximal *PI* minimal size	Closed maximal *PI* minimal range (interval)	Minimum size	Closed minimal range (interval)
De Swaan's data							
Mean number of successes (randomized scheme)	21.7889	29.6804	10.2334	4.2930	4.9607	4.6451	5.5605
Standard deviation (randomized scheme)	3.5623	3.1842	2.7641	1.9579	2.0839	2.0370	2.1892
Number of consistent predictions	54	43	39	27	30	26	34
Standardized deviation from randomized mean	9.0421	4.1831	10.4071	11.5975	12.0157	10.4836	12.9910
Israel (1949–2006 data)							
Mean number of successes (randomized scheme)	0.2947	7.6410	0.1567	0.0166	0.0200	0.4013	0.0195
Number of consistent predictions	6	11	4	0	1	1	2
$P_j(x \geq k_j^*)$	0.00000014	0.050929	0.000008818	1.0000	0.019855	0.3346	0.000159

requirement that the political parties are ranked accordingly. But this additional burden on the investigator appears to be amply compensated for by the very decent performance of two of the variations, especially by the *Closed Maximal PI* hypothesis and the *Closed Maximal PI Minimal Range* hypothesis.

When the *Closed Maximal PI* hypothesis is confronted with de Swaan's data set, it holds up well, achieving 39 successes out of 65 and a performance of 10.407 standard deviations above its randomized mean. It also performed exceedingly well in our supplementary analysis of the 18 Israeli elections and is probably the best performing of all hypotheses considered for this second data set, achieving 4 consistent predictions out of 18. We note that under the corresponding randomized scheme, the probability of observing 4 or more consistent predictions is only 0.000008 or practically zero. This is indicative that the level of predictive success achieved by the hypothesis is almost surely not the result of pure chance.

Equally encouraging results were obtained under the *Closed Maximal PI Minimal Range* hypothesis. This hypothesis which focuses attention only on those closed maximal *PI* coalitions that are of minimal range—worked well for the two data sets, producing predictions that are on par with those under the Leiserson-Axelrod closed minimal range theory, the latter being probably the best known predictive theory currently. The corresponding results for the Leiserson-Axelrod theory are reported in Table 1 for the purpose of comparison.

Under this variation of our alternative hypothesis, the level of predictive success obtained using the de Swaan data set is 30 out of 65 elections and this compares favorably with a predictive success rate of 34 out of 65 elections under the Leiserson-Axelrod closed minimal range theory; and in the case of the Israeli 1949–2006 data set, the relative performance of the two theories is one out of 18 and two out of 18 consistent predictions, respectively. It should be noted that except for one specific instance, this variation of the maximal *PI* hypothesis leads to predicted sets that are no larger than those under the Leiserson-Axelrod closed minimal range theory. In fact, in a number of instances, the predicted set under this variation is marginally smaller.

The variation that we have referred to as the *Closed Maximal PI Minimum Size* hypothesis, as expected, makes relatively precise predictions about the possible coalitions that will form in each situation but without the inherent instability property associated with the minimum size principle. This variation worked well under the de Swaan data set but did not produce a single consistent prediction in respect of the 18 Israeli elections thus leading us to, as in the case of the minimum size principle, reject the hypothesis.

5 Concluding Remarks

So, how useful are a priori power indices in predicting the formation of political coalitions? Do theories based on these indices work?

From our empirical investigation of Aumann's hypothesis (cf. Chua and Felsenthal 2008), which emphasizes a central role for the Shapley-Shubik index in the theory of coalition formation, we obtained an outcome that was largely negative. The prognosis flowing from the Felsenthal and Machover (2008) re-analysis of 77 alliances investigated in our test of Aumann's hypothesis is, at first blush, also not particularly promising. In 49 of the 77 alliances re-examined, the coalition leader is a posteriori a dictator, indicating in each of these 49 instances that the other members of the governmental alliance are a posteriori dummies. Furthermore, in 21 additional alliances, at least one member lost power or became a dummy after joining the alliance, raising the possibility that voting-power considerations leading to the formation of what Felsenthal and Machover (2002, 2008) called 'feasible' or 'expedient' alliances play no significant role in the formation of actual governmental alliances.[20]

But, from the perspective of the coalition leader, these observations are perhaps not surprising and even indicative that a priori measures of power do, in fact, play a role in the formation of governmental coalitions. Quite naturally, one would anticipate that a coalition leader would have an incentive to maneuver, as far as is possible, the coalitional arrangement to one in which s/he is a dictator within the alliance in the technical sense of the word. It would indeed be surprising if the opposite were true, that is, the coalition leader is a posteriori a dummy. Where the use of power indices fails is in their ability to explicitly capture the compromise that the coalition leader made in securing the a posteriori dictatorial or near dictatorial role. This, perhaps, may be a reason why, in our present analysis, the *Closed Maximal PI* and the *Closed Maximal PI Minimal Range* hypotheses performed relatively well; because the additional restriction(s) may have implicitly, perhaps only partially, captured the extent of the compromise that the coalition leader is required to make.

Viewed from this perspective, the measure of success achieved by the Penrose measure as reported in the present chapter suggests that further work should be carried out to more carefully investigate the role played by voting-power considerations in the formation of political coalitions.

Acknowledgments The authors wish to thank Matthew Braham, Moshé Machover, and Frank Steffen for their helpful comments. While working on this chapter Dan Felsenthal was also co-director of the Voting Power and Procedures Programme at the Centre for Philosophy of Natural and Social Science, London School of Economics and Political Science. This programme is supported by Voting Power in Practice Grant F/07 004/AJ from the Leverhulme Trust.

A.1 6 Appendix

Annex Tables 2, 3, 4, 5

[20] Felsenthal and Machover (2002, 2008) defined a *feasible alliance* as one in which the overall absolute voting power of each member of the alliance is not smaller than his voting power when no alliance exists. An *expedient alliance* is defined as one in which the overall absolute voting power of each member of the alliance is larger than his voting power when no alliance exists.

Annex Table 2 Summary of predictive accuracy (de Swaan's data)

No.	Country	Number of elections	Closed	Maximal *PI*	Closed maximal *PI*	Closed maximal *PI* minimal Size	Closed maximal *PI* minimal range (interval)	Minimum size	Closed minimal range (interval)
1	Denmark	17	16/75	16/98	15/40	15/18	15/20	16/18	16/22
2	Finland	10	7/65	2/49	2/19	1/10	1/10	1/13	2/13
3	France	1	1/9	0/3	0/1	0/1	0/1	0/1	0/2
4	Israel	6	4/83	6/449	4/53	1/6	1/10	0/13	1/10
5	Italy	2	2/20	1/55	1/12	1/2	1/4	1/2	1/4
6	Norway	4	3/18	4/37	3/11	3/4	3/5	4/4	3/7
7	Sweden	7	5/38	6/63	5/24	4/7	5/9	3/7	5/9
8	The Netherlands	15	14/171	6/199	7/38	2/15	4/19	1/29	6/21
9	Weimar Republic	3	2/31	2/84	2/7	0/3	0/4	0/5	0/5
	Total	65	54/510	43/1,037	39/205	27/66	30/82	26/92	34/93
	Mean Number of Successes (Randomized Scheme)	*	21.7889	29.6804	10.2334	4.2930	4.9607	4.6451	5.5605
	Standard Deviation (Randomized Scheme)	*	3.5623	3.1842	2.7641	1.9579	2.0839	2.0370	2.1892
	Standardized Deviation from Randomized Mean	*	9.0421	4.1831	10.4071	11.5975	12.0157	10.4836	12.9910

Annex Table 3 Detailed tabulation of predictive performance by country

No.	Election year	# winning coalitions	Closed	Maximal *PI*	Closed maximal *PI*	Closed maximal *PI* minimal size	Closed maximal *PI* minimal range (interval)	Minimum size	Closed minimal range (interval)
a. Denmark									
1	1918	6	1/5	0/4	0/3	0/1	0/2	1/1	1/2
2	1920	6	1/4	1/3	1/2	1/1	1/1	1/1	1/1
3	1920A	6	1/4	1/3	1/2	1/1	1/1	1/1	1/1
4	1920B	6	1/4	1/3	1/2	1/1	1/1	1/1	1/1
5	1924	7	1/3	1/4	1/1	1/1	1/1	1/1	1/1
6	1926	6	1/4	1/2	1/2	1/1	1/1	1/1	1/1
7	1929	7	1/3	1/5	1/2	1/1	1/1	1/1	1/1
8	1932	7	1/3	1/5	1/2	1/1	1/1	1/1	1/1
9	1935	7	1/3	1/6	1/2	1/1	1/1	1/1	1/1
10	1939	7	1/3	1/6	1/2	1/1	1/1	1/1	1/1
11	1945	11	1/5	1/1	1/1	1/1	1/1	1/1	1/2
12	1953	14	1/6	1/11	1/4	1/1	1/1	1/1	1/1
13	1957	13	0/3	1/9	0/1	0/1	0/1	1/1	0/1
14	1960	15	1/7	1/12	1/5	1/2	1/2	1/2	1/2
15	1966	14	1/6	1/11	1/4	1/1	1/1	1/1	1/1
16	1968	10	1/5	1/3	1/1	1/1	1/1	0/1	1/2
17	1971	15	1/7	1/10	1/4	1/1	1/2	1/1	1/2
	Total	157	16/75	16/98	15/40	15/18	15/20	16/18	16/22
b. Finland									
1	1924	20	1/3	0/6	1/2	1/1	1/1	0/1	1/1
2	1930	12	1/2	0/5	0/1	0/1	0/1	0/1	1/1
3	1933	17	1/6	0/2	0/1	0/1	0/1	1/2	0/1
4	1939	16	1/6	0/1	0/1	0/1	0/1	0/1	0/1
5	1951	24	0/9	0/2	0/1	0/1	0/1	0/3	0/1

(continued)

Annex Table 3 (continued)

No.	Election year	# winning coalitions	Closed	Maximal *PI*	Closed maximal *PI*	Closed maximal *PI* minimal size	Closed maximal *PI* minimal range (interval)	Minimum size	Closed minimal range (interval)
b. Finland (cont'd)									
6	1954	18	0/5	1/3	0/2	0/1	0/1	0/1	0/2
7	1958	43	0/8	0/1	0/4	0/1	0/1	0/1	0/1
8	1962	23	1/8	0/1	0/1	0/1	0/1	0/1	0/2
9	1966	49	1/9	0/6	0/1	0/1	0/1	0/1	0/2
10	1970	42	1/9	1/22	1/5	0/1	0/1	0/1	0/1
	Total	264	7/65	2/49	2/19	1/10	1/10	1/13	2/13
c. France									
1	1947	44	1/9	0/3	0/1	0/1	0/1	0/1	0/2
	Total	44	1/9	0/3	0/1	0/1	0/1	0/1	0/2
d. Israel									
1	1949	115	1/16	1/83	1/10	0/1	0/1	0/1	0/1
2	1955	226	1/19	1/112	1/9	0/1	0/1	0/6	0/1
3	1959	118	1/15	1/98	1/11	0/1	0/3	0/2	0/3
4	1961	112	1/15	1/61	1/8	1/1	1/1	0/1	1/1
5	1965	54	0/9	1/36	0/7	0/1	0/2	0/2	0/2
6	1969	60	0/9	1/59	0/8	0/1	0/2	0/1	0/2
	Total	685	4/83	6/449	4/53	1/6	1/10	0/13	1/10
e. Italy									
1	1946	52	1/11	0/33	0/6	0/1	0/1	0/1	0/1
2	1972	29	1/9	1/22	1/6	1/1	1/3	1/1	1/3
	Total	81	2/20	1/55	1/12	1/2	1/4	1/2	1/4
f. Norway									
1	1936	7	0/3	1/6	0/2	0/1	0/1	1/1	0/1
2	1961	31	1/9	1/29	1/7	1/1	1/2	1/1	1/2

(continued)

Annex Table 3 (continued)

No.	Election year	# winning coalitions	Closed	Maximal *PI*	Closed maximal *PI*	Closed maximal *PI* minimal size	Closed maximal *PI* minimal range (interval)	Minimum size	Closed minimal range (interval)
3	1965	9	1/3	1/1	1/1	1/1	1/1	1/1	1/2
4	1969	9	1/3	1/1	1/1	1/1	1/1	1/1	1/2
	Total	56	3/18	4/37	3/11	3/4	3/5	4/4	3/7
g. Sweden									
1	1917	12	0/6	0/1	0/1	0/1	0/1	0/1	0/1
2	1924	7	0/3	1/6	0/2	0/1	0/1	0/1	0/1
3	1932	14	1/6	1/12	1/4	1/1	1/1	0/1	1/1
4	1936	15	1/7	1/14	1/6	0/1	1/2	0/1	1/2
5	1952	7	1/3	1/6	1/2	1/1	1/1	1/1	1/1
6	1956	14	1/6	1/11	1/4	1/1	1/1	1/1	1/1
7	1970	15	1/7	1/13	1/5	1/1	1/2	1/1	1/2
	Total	84	5/38	6/63	5/24	4/7	5/9	3/7	5/9
h. The Netherlands									
1	1918	95	1/14	1/30	1/5	1/1	1/2	1/3	1/2
2	1922	25	1/9	1/10	1/3	0/1	1/3	0/1	1/3
3	1925	19	1/8	0/1	0/1	0/1	0/1	0/1	1/2
4	1929	48	1/11	1/11	1/1	1/1	1/1	0/1	1/2
5	1933	79	1/14	0/6	0/1	0/1	0/1	0/1	0/1
6	1937	83	1/15	0/9	0/1	0/1	0/1	0/3	1/2
7	1946	24	1/10	1/8	1/3	0/1	1/1	0/1	1/1
8	1948	21	0/8	0/1	0/2	0/1	0/2	0/1	0/1
9	1952	23	1/8	0/5	1/8	0/1	0/1	0/2	0/1
10	1956	24	1/8	0/9	0/1	0/1	0/1	0/1	0/1
11	1959	12	1/6	1/5	1/3	0/1	0/1	0/1	0/1
12	1963	48	1/14	1/18	1/4	0/1	0/1	0/1	0/1

(continued)

Annex Table 3 (continued)

No.	Election year	# winning coalitions	Closed	Maximal *PI*	Closed maximal *PI*	Closed maximal *PI* minimal size	Closed maximal *PI* minimal range (interval)	Minimum size	Closed minimal range (interval)
h. The Netherlands (cont'd)									
13	1967	192	1/20	0/29	0/1	0/1	0/1	0/4	0/1
14	1971	64	1/11	0/3	0/2	0/1	0/1	0/2	0/1
15	1973	202	1/15	0/54	0/2	0/1	0/1	0/6	0/1
	Total	959	14/171	6/199	7/38	2/15	4/19	1/29	6/21
i. Weimar Republic									
1	1919	56	1/10	1/39	1/3	0/1	0/1	0/1	0/1
2	1925	150	0/12	0/4	0/1	0/1	0/1	0/2	0/2
3	1928	105	1/9	1/41	1/3	0/1	0/2	0/2	0/2
	Total	311	2/31	2/84	2/7	0/3	0/4	0/5	0/5

Annex Table 4 Summary of predictive accuracy (Israel 1949–2006)

No.	Election year	# winning coalitions	Closed	Maximal *PI*	Closed maximal *PI*	Closed maximal *PI* minimal size	Closed maximal *PI* minimal range (interval)	Minimum size	Closed minimal range (interval)
1	1949	1,883	0/30	1/1,258	0/15	0/2	0/1	0/18	0/1
2	1951	15,262	0/42	1/9,951	0/32	0/1	0/1	0/193	0/1
3	1955	1,867	1/30	0/816	0/14	0/1	0/1	0/24	0/1
4	1959	1,961	1/35	1/1,468	1/25	0/1	0/1	0/16	0/1
5	1961	929	1/25	1/477	1/15	0/1	0/4	0/17	0/4
6	1965	3,733	0/41	1/2,220	0/31	0/1	0/1	0/60	0/1
7	1969	4,069	0/44	1/3,905	0/39	0/2	0/2	0/18	0/2
8	1973	454	0/25	1/282	0/19	0/1	0/1	0/13	0/1
9	1977	3,470	0/14	1/1,738	0/4	0/1	0/1	0/58	0/1
10	1981	384	1/13	1/129	1/5	0/2	0/2	0/20	0/2
11	1984	12,875	0/27	0/4,693	0/12	0/2	0/2	0/587	0/1
12	1988	12,288	0/30	0/4,094	0/2	0/1	0/1	0/492	0/1
13	1992	440	1/22	1/216	1/14	0/1	1/1	0/10	1/1
14	1996	739	1/12	0/57	0/1	0/1	0/1	0/20	0/1
15	1999	12,703	0/39	0/2	0/1	0/1	0/1	0/301	0/3
16	2001	11,059	0/23	0/3	0/2	0/1	0/2	0/297	0/2
17	2003	3,618	0/16	1/1,218	0/2	0/1	0/1	0/57	0/1
18	2006	1,641	0/33	0/11	0/2	0/1	0/2	1/33	1/1
Total		89,375	6/501	11/32,538	4/235	0/22	1/26	1/2,234	2/26
Mean number of successes (randomized scheme)		*	0.2947	7.6410	0.1567	0.0166	0.0200	0.4013	0.0195
Number of consistent predictions		*	6	11	4	0	1	1	2
$P_j(x \geq k_j^*)$		*	0.00000014	0.050929	0.000008818	1.0000	0.019855	0.3346	0.000159

Annex Table 5 Exact probability mass function $P_j(x = k)$ for randomized scheme under each theory

Number of successes (k)	$^{n}C_{k}$	Closed	Maximal *PI*	Closed maximal *PI*	Closed maximal *PI* minimal size	Closed maximal *PI* minimal range (interval)	Minimum size	Closed minimal range (interval)
0	1	0.74122	2.14E–06	0.8534	0.98349	0.98015	0.66543	0.98067
1	18	0.22557	8.51E–05	0.13691	0.016393	0.019686	0.27515	0.019176
2	153	0.030609	0.0010566	0.0093215	0.00011546	0.000168	0.052672	0.000158
3	816	0.00246	0.006855	0.00035783	4.63E–07	8.07E–07	0.0062037	7.31E–07
4	3,060	0.000131	0.027705	8.68E–06	1.19E–09	2.45E–09	0.00050386	2.12E–09
5	8,568	4.92E–06	0.075906	1.41E–07	2.10E–12	5.02E–12	3.00E–05	4.15E–12
6	18,564	1.35E–07	0.14763	1.57E–09	2.64E–15	7.19E–15	1.35E–06	5.68E–15
7	31,824	2.74E–09	0.20905	1.22E–11	2.42E–18	7.40E–18	4.74E–08	5.60E–18
8	43,758	4.20E–11	0.21814	6.65E–14	1.64E–21	5.57E–21	1.31E–09	4.07E–21
9	48,620	4.87E–13	0.168	2.51E–16	8.32E–25	3.10E–24	2.85E–11	2.21E–24
10	43,758	4.28E–15	0.094646	6.50E–19	3.16E–28	1.29E–27	4.95E–13	8.99E–28
11	31,824	2.83E–17	0.0382	1.15E–21	8.96E–32	3.97E–31	6.79E–15	2.74E–31
12	18,564	1.40E–19	0.010627	1.37E–24	1.88E–35	9.05E–35	7.32E–17	6.25E–35
13	8,568	5.13E–22	0.0019021	1.08E–27	2.88E–39	1.51E–38	6.11E–19	1.05E–38
14	3,060	1.35E–24	0.00019178	5.48E–31	3.15E–43	1.80E–42	3.85E–21	1.26E–42
15	816	2.50E–27	8.00E–06	1.68E–34	2.37E–47	1.48E–46	1.77E–23	1.05E–46
16	153	3.05E–30	4.87E–08	2.90E–38	1.16E–51	7.89E–51	5.56E–26	5.72E–51
17	18	2.21E–33	1.98E–11	2.50E–42	3.31E–56	2.45E–55	1.06E–28	1.80E–55
18	1	7.13E–37	1.93E–15	8.08E–47	4.15E–61	3.32E–60	9.31E–32	2.49E–60
Sum	65,536	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

References

Axelrod, R. (1970). *Conflict of interest: A theory of divergent goals with applications to politics*. Chicago: Markham.

Chua, V. C. H., & Felsenthal D. S. (2008). Coalition formation theories revisited: An empirical investigation of Aumann's hypothesis. In M. Braham & F. Steffen (Eds.), *Power, freedom, and voting: Essays in honour of Manfred J. Holler*. Berlin Heidelberg: Springer.

Coleman, J. S. (1971). Control of collectivities and the power of a collectivity to act. In B. Lieberman (Ed.), *Social choice*. New York: Gordon and Breach.

de Swaan, A. (1973). *Coalition theories and cabinet formation: A study of formal theories of coalition formation applied to nine European parliaments after 1918*. Amsterdam: Elsevier Scientific Publishing Company.

Felsenthal, D. S., & Machover, M. (1998). *The measurement of voting power*. Cheltenham: Edward Elgar.

Felsenthal, D. S., & Machover, M. (2002). Annexations and alliances: When are blocs advantageous a priori? *Social Choice and Welfare, 19*, 295–312.

Felsenthal, D. S., & Machover, M. (2004). A priori voting power: What is it all about? *Political Studies Review, 2*, 1–23.

Felsenthal, D. S. & Machover M. (2008). Further reflections on the expediency and stability of alliances. In M. Braham & F. Steffen (Eds.), *Power, freedom, and voting: Essays in honour of Manfred J. Holler*. Berlin Heidelberg: Springer.

Gamson, W. A. (1961). A theory of coalition formation. *American Sociological Review, 26*, 373–382.

Laver, M., & Schofield, N. (1990). *Multiparty government: Politics of coalitions in Europe*. Oxford: Oxford University Press.

Leiserson, M. A. (1966). Coalitions in politics: A theoretical and empirical study. *Mimeographed doctoral dissertation*, Yale University, New Haven.

Penrose, L. S. (1952). *On the objective study of crowd behavior*. London: H.K. Lewis.

Riker, W. H. (1959). A test of the adequacy of the power index. *Behavioral Science, 4*, 120–131.

Riker, W. H. (1962). *The theory of political coalitions*. New Haven: Yale University Press.

Ross, S. M. (2002). *A first course in probability theory* (6th ed.). Upper Saddle River: Prentice Hall Publishing Company.

Shapley, L. S. (1953). A value for *n*-person games. In H. W. Kuhn & A. W. Tucker (Eds.), *Contributions to the theory of games II (Annals of Mathematics Studies)*. Princeton: Princeton University Press.

Shapley, L. S., & Shubik, M. J. (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review, 48*, 787–792.

van Damme, E. (1998). On the state of the art in game theory: An interview with Robert Aumann. *Games and Economic Behavior, 24*, 181–210.

A Note on Communication Structures

Vito Fragnelli

1 Introduction

In several decisional situations the result is simply in favor or against a proposal, with no intermediate position, so that they are often referred to as *0-1 decision problems*, where 1 means that the proposal is accepted and 0 that it is rejected. Accordingly, the decision-makers express their preferences to be favorable or not to the proposal. The most common of these situations take place in Parliamentary voting sessions, board of directors decisions and so on. An important question is how to evaluate the influence of each member on the final decision, especially when the members are not equivalent, for instance because they are political parties with different numbers of seats in the Parliament or stakeholders endowed with different stock shares. This analysis may be performed, inter alia, by using power indices.

A simple representation of decisional situations is through a *weighted majority situation* $[q; w_1, w_2, ..., w_n]$, where given a set $N = \{1, 2, ..., n\}$ of decision-makers, w_i is the weight of agent $i \in N$, e.g. the number of seats of a party or the stock share of a stakeholder, and q is the majority quota; the interpretation is that a subset $S \subseteq N$ of decision-makers is able to pass a proposal if and only if $\sum_{i \in S} w_i \geq q$. It is possible to associate a *weighted majority game* (N, v) defined as:

$$v(S) = \begin{cases} 1 & \text{if } \sum_{i \in S} w_i \geq q \\ 0 & \text{otherwise} \end{cases}, \quad S \subseteq N$$

A coalition $S \subseteq N$ is called *winning* if $v(S) = 1$ and *losing* if $v(S) = 0$.

This chapter has been published in Homo Oeconomicus 29(2), 2012.

V. Fragnelli (✉)
Department of Science and Technological Innovation, University of Eastern Piedmont,
Viale T. Michel 11, 15121 Alessandria, Italy
e-mail: vito.fragnelli@mfn.unipmn.it

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_24, © Springer-Verlag Berlin Heidelberg 2013

We start shortly surveying the literature on power indices. The large number of existing power indices depends on the matter that each emphasizes different features of the problem, making it particularly suitable for specific situations.

The first indices (see Penrose 1946; Shapley and Shubik 1954; Banzhaf 1965; Coleman 1971) are based on the ability of a decision-maker to switch the result of the voting from rejection to approval by joining a set of other decision-makers that are in favor of the proposal. More precisely, the indices of Penrose, Banzhaf and Coleman tally the switches w.r.t. the possible coalitions, while in the Shapley-Shubik index also the order agents form a coalition plays a role.

In 1977 two important contributions strongly modified the prevailing power indices, introducing the element of possible relations among the agents. Myerson (1977) proposed to use an undirected graph, called *communication structure*, in order to represent the relationships among the decision-makers and determining the Shapley value of a suitably restricted game. Owen (1977) introduced the *a priori unions*, or *coalition structures*, that account for existing agreements, not necessarily binding, among some decision-makers. The relationship among the power indices, mainly Shapley-Shubik and Banzhaf, in the original game and in the restricted game á la Myerson was studied in Owen (1986), with particular attention to those situations in which the underlying graph is a tree.

Recently, in Khmelnitskaya (2007) the communication structures by Myerson and the coalition structures by Owen were combined to define a new index.

In the following years, several new indices were introduced for better representation of other situations. Deegan and Packel (1978) defined a new index that considers only the *minimal winning coalitions*, i.e. those coalitions in which each agent is critical, in the sense that the coalition becomes losing when s/he leaves; Johnston (1978) used a similar model but he considered the *quasi-minimal winning coalitions*, i.e. those coalitions in which at least one agent is critical. Both indices are based on the idea of dividing the unitary power among the minimal or quasi-minimal winning coalitions, respectively; then the power assigned to each coalition is equally shared among the critical agents in it. We may say that they suppose that a very large coalition, in which no agent is critical, has no possibility to form, as the “cake should be divided among too many agents”.

Holler (1982) introduced the Public Good index, supposing that the worth of a coalition is a *public good*, so the members of the winning decisive sets, i.e. the minimal winning coalitions, have to enjoy the same value. The result is that the power of an agent is proportional to the number of minimal winning coalitions s/he belongs to.

Another important contribution was due to Kalai and Samet (1987). They introduced the idea of adding a *weight* to the elements characterizing each agent, besides her/his ability of switching the result; Haeringer (1999) combined the idea of the weights by Kalai and Samet with the idea of communication structure.

In 1989 Winter extended the idea of a priori unions by Owen by requiring that the different unions may join only according to a predefined scheme, that he called *levels structure*, and introduced the levels structure value.

In the following years, some papers dealt with situations in which only some coalitions are feasible, not only for communication reasons but also for possible incompatibilities among the agents; two main research fields are related to the permission structures and to special structures of the feasible coalitions. In the first group we mention the papers by Gilles et al. (1992), Van den Brink and Gilles (1996) and Van den Brink (1997). In the second group we cite the papers by Bilbao et al. (1998) and by Bilbao and Edelman (2000) that consider the Banzhaf index and the Shapley value on convex geometries, respectively and the papers by Algaba et al. (2003, 2004) that study the Shapley value and the Banzhaf value on antimatroids, respectively. We refer to the paper by Katsev (2010) for a survey on values for games with restricted cooperation.

Fragnelli et al. (2009) introduced a new family of power indices, called *FP*, that account the issue of *contiguity* in a monodimensional voting space; the indices in this family require to select a set of contiguous winning coalitions among which the unitary power is divided accounting the probability that each coalition form, then the power of each coalition is shared among its parties according to their relevance in the coalition. The idea of contiguity was extended to the idea of *connectedness* in a possibly multidimensional voting space by Chessa and Fragnelli (2011). In both cases, non-contiguous and non-connected coalitions are ignored. The idea of monodimensionality of the voting space was already considered in Amer and Carreras (2001).

We refer to the papers for the formal definitions of the indices and for further details.

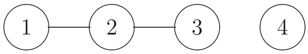
In this note, we concentrate on the concept of incompatible agents. Communication structures provide a very powerful tool for representing incompatibilities, but it is necessary to carefully analyze them in order to decide which structure is more suitable in a given majority situation, in which the agents are not available to form any (theoretically) possible coalition.

2 Communication Structures and Incompatible Agents

In Myerson (1977) considers a situation in which communication is represented by using an undirected graph whose vertices correspond to the agents and the edges connect pairs of agents that are compatible (or may communicate). In the *restricted game* introduced by Myerson, a coalition is feasible and its worth is “effective” if the vertices associated to its players are connected, otherwise the worth of the coalition is the sum of the worths of the subcoalitions in the partition in connected components induced by the communication graph.

It is possible to raise some questions about this approach in a weighted majority situation setting; for doing so, let us consider the following example.

Example 1 Let $N = \{1,2,3,4\}$ be the set of parties of the weighted majority situation $[51; 35, 30, 25, 10]$; the winning coalitions are $\{1,2\}$, $\{1,3\}$, $\{2,3\}$, $\{1,2,3\}$, $\{1,2,4\}$, $\{1,3,4\}$, $\{2,3,4\}$, $\{1,2,3,4\}$ and suppose that the communication structure is represented by the graph G in the following figure:



Player 4 is not connected and therefore will not be in a feasible coalition with one of the other players; also players 1 and 3 are not directly connected. In the restricted game (N, v_G) induced by the graph G, the winning coalitions reduce to $\{1,2\}$, $\{2,3\}$, $\{1,2,3\}$, $\{1,2,4\}$, $\{2,3,4\}$, $\{1,2,3,4\}$; for instance, $v_G(\{1,3\}) = v(\{1\}) + v(\{3\}) = 0$ while $v_G(\{1,2,4\}) = v(\{1,2\}) + v(\{4\}) = 1$.

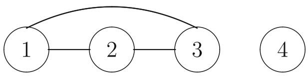
We may provide some comments.

(i) The games v and v_G are monotonic; in fact, $v_G(\{1,2\}) = 1$ and $v_G(\{1,2,4\}) = 1$, even if the second coalition is not feasible, being not connected.

(ii) Dealing with infeasible coalitions, we may have some problems for computing indices based on the marginal contributions or swings. Let us consider the weighted majority situation $[51; 24, 25, 51]$ in which coalitions have to be contiguous, so the feasible winning coalitions are $\{3\}$, $\{2,3\}$, $\{1,2,3\}$ and party 3 should have all the power. On the other hand the infeasible coalition $\{1,3\}$ needs some comments. If we assign it value 0, we face the situation in which party 1 enters the winning coalition $\{3\}$ generating the losing coalition $\{1,3\}$, i.e. party 1 has a negative marginal contribution or a negative swing; moreover, party 2 results to be critical for coalition $\{1,2,3\}$, so it has a positive marginal contribution or a positive swing. A possible solution is to account only the marginal contributions and the swings that involve pairs of feasible coalitions, including losing ones, and to modify the definition of the indices accordingly. Of course, this problem does not show up with those indices, like Deegan-Packel, Johnston and Holler, which do not compare pairs of coalitions.

(iii) In Example 1, the graph G indicates that coalitions $\{1,2\}$, $\{2,3\}$, $\{1,2,3\}$ are feasible while coalition $\{1,3\}$ is infeasible. Now, let us suppose that parties 1 and 3 are both available for forming a two-party majority with party 2 but they never want to stay in the same coalition, so that also coalition $\{1,2,3\}$ is infeasible; the previous approach does not enable us to represent this situation.

(iv) Referring to the previous point, we can modify the definition of feasibility saying that a coalition is feasible if the corresponding subgraph is complete (*clique*). In this way the graph G represents the situation in which coalitions $\{1,2\}$, $\{2,3\}$ are feasible and coalitions $\{1,3\}$, $\{1,2,3\}$ are infeasible. If we want to represent a situation in which coalition $\{1,2,3\}$ is feasible, we may consider the graph G', where an edge between parties 1 and 3 is added:



Unfortunately, the new graph G' indicates that parties 1 and 3 accept forming a majority also without party 2. The problem cannot be solved neither introducing oriented arcs instead of unoriented edges.

(v) Supposing that a suitable tool for representing all the feasible coalitions, and no more, is available, all these coalitions are usually considered equivalent, so another interesting question arises: how to evaluate the probability that a coalition forms? Having in mind the work by Calvo et al. (1999), we may associate to edge $(i,j), i,j \in N$ a real number $0 \leq p_{ij} \leq 1$ that can be viewed as the probability that the two parties i and j enter the same coalition, so we have to find a method for computing the probability p_S of each feasible coalition $S \subseteq N$.

A simple idea is to define p_S as the product of the weights of the edges of the subgraph associated to S. We may remark that this method requires assuming that the events that two parties join are independent. Then, we may observe that $S \subset T$ implies $p_S \geq p_T$; if we account only minimal winning coalitions, T is not minimal and then it is excluded. In the other situations, it is equivalent to the hypothesis that a larger coalition has a larger cost so its probability decreases. But also in this case, we may raise a question. For instance, if we compute $p_{\{1,2,3\}}$ referring to the graph G it is given by $p_{12}p_{23}$ that is greater than or equal to $p_{12}p_{23}p_{13}$ that is the value obtained referring to the graph G'. On the other hand, it is reasonable to expect a larger probability when we refer to the graph G' where the three parties have stronger connections. This situation may be solved by imposing a complete graph, i.e. all pairs of nodes are connected by an arc. This approach was introduced by Calvo et al. (1999) for introducing the *probabilistic extension* of the Myerson value. They suppose that all pairs of agents may join according to a probability of direct communication that can be viewed as a degree of cooperation. The probability that a coalition of agents forms when a probabilistic communication graph is assigned is the product of the probabilities of the edges in the graph times the probability that the other edges of the subgraph associated to the coalition are not used. Finally, they define a restricted game in which the value of each coalition is the sum over all the subgraphs defined on the vertices associated to the coalition itself of the product of the probabilities that the coalition forms with each communication subgraph times the value of the coalition in the restricted game á la Myerson in that subgraph. Another approach is to define the probability of a coalition as the sum of the probabilities of the edges of the corresponding subgraph. Again, we may remark that it corresponds to require that the events that two parties enter the same coalition are disjoint, otherwise the resulting probability could be larger than 1. Finally, we remark that the probability to form for the different coalitions cannot be based on the same hypotheses whatever the parties entering the coalitions and that the parties in a coalition may influence the probability of the entrance of a new party. For instance, a left party will enter a right parties coalition with a low probability, but the probability increases if the coalition includes center parties and increases more if the resulting coalition is contiguous.

3 Concluding Remarks

Summarizing, we may say that, in our opinion, a good way to deal with incompatible agents is to use a graph to represent compatibilities among pairs of agents; then a set of feasible coalitions (or "relevant" coalitions) can be identified as a subset of all the connected coalitions; finally, a probability is assigned to each feasible coalition, referring to opinions of a panel of experts, or to the majorities formed in the past. This approach does not require the estimation of the probabilities of all the pairs of agents (for instance, see Remark 5.3 in Calvo et al. 1999, and does not need any hypothesis for adding or multiplying the probabilities of the pairs for assigning the probabilities of the coalitions.

We may also notice that this idea fits very well the requirements of the *FP* indices for contiguous/connected coalitions (see Fragnelli et al. 2009; Chessa and Fragnelli 2011) we can define the weights of the coalitions as their normalized probabilities, and the weights of the parties in each coalition as their percentages of seats in that coalition.

As we said in the previous section, another interesting open problem is the modification of indices based on marginal contributions and swings when some coalitions are infeasible.

Acknowledgments The author gratefully acknowledges Manfred Holler for his valuable suggestions.

References

Algaba, E., Bilbao, J. M., Van den Brink, R., & Jiménez-Losada, A. (2003). Axiomatizations of the restricted Shapley value for cooperative games on antimatroids. *Mathematical Methods of Operations Research*, *57*, 49–65.

Algaba, E., Bilbao, J. M., Van den Brink, R., & Jiménez-Losada, A. (2004). An axiomatization of the Banzhaf value for cooperative games on antimatroids. *Mathematical Methods of Operations Research*, *59*, 147–166.

Amer, R., & Carreras, F. (2001). Power, cooperation indices and coalition structures. In M. J. Holler & G. Owen (Eds.), *Power indices and coalition formation* (pp. 153–173). Dordretch: Kluwer Academic Publishers.

Banzhaf, J. F. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review*, *19*, 317–343.

Bilbao, J. M., Jiménez, A., & Lopez, J. J. (1998). The Banzhaf power index on convex geometries. *Mathematical Social Sciences*, *36*, 157–173.

Bilbao, J. M., & Edelman, P. H. (2000). The Shapley value on convex geometries. *Discrete Applied Mathematics*, *103*, 33–40.
Calvo, E., Lasaga, J., & Van den Nouweland, A. (1999). Values of games with probabilistic graphs. *Mathematical Social Sciences*, *37*, 79–95.
Chessa, M., & Fragnelli, V. (2011). Embedding classical indices in the FP family. *AUCO Czech Economic Review*, *5*, 289–305.
Coleman, J. S. (1971). Control of collectivities and the power of a collectivity to act. In B. Lieberman (Ed.), *Social Choice* (pp. 269–300). London: Gordon and Breach.
Deegan, J., & Packel, E. W. (1978). A new index of power for simple n-person games. *International Journal of Game Theory*, *7*, 113–123.
Fragnelli, V., Ottone, S., & Sattanino, R. (2009). A new family of power indices for voting games. *Homo Oeconomicus*, *26*, 381–394.
Gilles, R. P., Owen, G., & Van den Brink, R. (1992). Games with permission structures: The conjunctive approach. *International Journal of Game Theory*, *20*, 277–293.
Haeringer, G. (1999). Weighted Myerson value. *International Game Theory Review*, *1*, 187–192.
Holler, M. J. (1982). Forming coalitions and measuring voting power. *Political Studies*, *30*, 262–271.
Johnston, R. J. (1978). On the measurement of power: Some reactions to Laver. *Environment and Planning A*, *10*, 907–914.
Kalai, E., & Samet, D. (1987). On weighted Shapley values. *International Journal of Game Theory*, *16*, 205–222.
Katsev, I.V. (2010). The Shapley value for games with restricted cooperation. Available at SSRN: http://ssrn.com/abstract=1709008
Khmelnitskaya, A. B. (2007). *Values for graph-restricted games with coalition structure, memorandum 1848*. Enschede: University of Twente.
Myerson, R. (1977). Graphs and cooperation in games. *Mathematics of Operations Research*, *2*, 225–229.
Owen, G. (1977). Values of games with a priori unions. *Lecture Notes in Economic and Mathematical Systems*, *141*, 76–88.
Owen, G. (1986). Values of graph-restricted games. *SIAM Journal of Algebraic Discrete Methods*, *7*, 210–220.
Penrose, L. S. (1946). The elementary statistics of majority voting. *Journal of the Royal Statistical Society*, *109*, 53–57.
Shapley, L. S., & Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review*, *48*, 787–792.
Van den Brink, R., & Gilles, R. P. (1996). Axiomatizations of the conjunctive permission value for games with permission structures. *Games and Economic Behavior*, *12*, 113–126.
Van den Brink, R. (1997). An axiomatization of the disjunctive permission value for games with a permission structure. *International Journal of Game Theory*, *26*, 27–43.
Winter, E. (1989). A value for cooperative games with levels structure of cooperation. *International Journal of Game Theory*, *18*, 227–240.

Shareholder Voting Power and Ownership Control of Companies

Dennis Leech

1 Introduction

In countries such as the UK and USA, where the ownership of industry is widely dispersed among a large number of shareholders, with few companies having a majority shareholder, shareholder power is a matter of academic debate. The literature that discusses their role in corporate governance tends to divide sharply between two extremes represented, on one side, mainly by academic economists (for example Shleifer and Vishny (1997)), who emphasise the effect of ownership dispersion on the incentives of managers. These writers sidestep all questions surrounding the fundamental nature and direction of the firm, disallowing any role for ownership in control. Their view of the firm is of a kind of machine for making profits, the important question being whether its managers choose to make more or less; all decisions are more or less clear cut and corporate governance is a matter of getting the incentives right. On the other side is a smaller, less theoretical literature, deriving largely from business schools and independent corporate governance consultants, that stresses voting by shareholders and the influence deriving from it. Writers in this tradition tend to regard a firm, specifically one listed on the stock

An earlier version of this chapter has been published in *Power Measures*, Vol II (*Homo Oeconomicus* 19(3)), edited by Manfred J. Holler and Guillermo Owen, Munich: Accedo Verlag, 2002: 345–371. The ideas and results in this chapter have been presented in different forms at the conferences of the European Association for Research in Industrial Economics in Turin, September 1999, "Ownership and Economic Performance", Oslo May 2000 and the Game Theory Society, Games 2000, Bilbao, July 2000, and to staff seminars at City, Oxford, Oslo, Warwick and Cambridge Universities.

D. Leech (✉)
Department of Economics, University of Warwick, Coventry, CV4 7AL, UK
e-mail: d.leech@warwick.ac.uk
www.warwick.ac.uk/fac/soc/Economics/leech

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*, DOI: 10.1007/978-3-642-35929-3_25, © Springer-Verlag Berlin Heidelberg 2013

exchange whose shares are openly traded, as a public institution more broadly defined and therefore corporate governance deals with a greater range of concerns than those cast in narrowly economic terms.

The dominant view on shareholder power among the first group, which derives from Berle and Means (1932), is that because of ownership dispersion, shareholders as a group are powerless because no individual among them could be said to have any appreciable voting power or control. Using ownership dispersion as a paradigm, they have emphasised the moral hazard argument that in general even a relatively large shareholder has little incentive to monitor the performance of the management, to take an interest in the direction of the firm or even to vote their shares because their ownership stake—small in percentage terms—gives them only a small entitlement to the returns accruing to their investment in those activities. The second group tends to advocate better standards of corporate governance through greater shareholder involvement. They base their position on the power of the vote as a fundamental part of equity ownership and have a different paradigm: that of a world of large institutional investors whose power and influence derives from managing huge pension and insurance funds. A firm is often faced with fundamental uncertainty about its nature and direction so that strategic decisions must be taken in the absence of full information. Shareholders have the central responsibility for all this; there is considerably more to their role than simply designing mechanisms to motivate managers to maximise the value of the firm. The role of shareholders is to determine all questions that are not routine, that cannot be decided by management because they are relevant to the relationship between the company and the capital market, and these include all fundamental matters affecting it.

This chapter is a contribution to the literature on the latter question. I abstract from the question of incentives here and concentrate solely on voting power. I also apply this to study the idea of minority control first proposed by Berle and Means. For present purposes I maintain the assumption that shareholders always do have incentives to take part in monitoring the management and voting their shares. This is obviously an unreal assumption because it is not worthwhile for very small shareholders with holdings of only a few thousand pounds to be active owners, but it biases the analysis away from finding shareholders to be powerful. Therefore the results of the analysis are stronger to the extent they point to evidence of shareholders as being powerful. If the analysis is confined to larger shareholders who have incentives to be active deriving from the size of their holding, then it is likely to find some of them to be very powerful. However that depends on having a model of shareholder incentives and is left to another study. I am here exclusively concerned with the important conceptual and statistical issue of the relation between the size of an ownership stake and the power or control it represents.

2 Corporate Governance and Shareholders

That the term corporate governance has come to be used for the system by which firms are regulated is a testimony to shareholder power in both the United States and the United Kingdom. Corporate governance became a policy focus in the United States earlier as a result of the growth of financial institutions especially public sector pension funds who found themselves to be relatively important shareholders and saw that they had a direct interest in being active as owners. In the 1980s they became concerned to protect their rights as owners against attempts by the top management of many companies to introduce anti-takeover measures and effectively limit their accountability to shareholders. Their resulting campaign preserve and strengthen shareholder democracy was successful and they later moved on to using their voting power as a weapon to improve performance. Some financial institutions, for example Calpers, the California Public Employees pension scheme, explicitly adopted a policy of engagement with management of underperforming companies, threatening to use their voting power to force changes in the board and replace the chief executive if performance targets were not met. A factor that contributed to this change of approach—institutional shareholders had previously restricted themselves to a role as active managers of a portfolio, reacting to underperformance by a company by selling its shares—was the large size of some of the holdings which made it difficult to sell without damaging the market. Also many of them held essentially passive portfolios. Thus the policy developed, articulated by Calpers, was one of "active ownership combined with passive investment"[1] which depended on good standards of corporate governance in terms of the relationships between firms and shareholders.

The corporate governance movement in the UK originated differently and somewhat later. In the early eighties a number of rather high profile corporate failures (Maxwell, Polly Peck, BCCI) that were due to mismanagement or wrongdoing by management showed fundamental inadequacies in the system of corporate governance. Investors themselves rather than the government decided that it was necessary to act and so the response was to introduce a voluntary code of practice, rather than government legislation, to ensure high standards of behaviour. The fundamental principle behind this approach was shareholder power based on greater disclosure of information and also structural changes in the way boards are run leading to greater accountability of management, and shareholders being expected to discharge responsibilities as owners by actively engaging with management and voting their shares. The Cadbury Code, recommended a number of changes in the way boards operate, such as the separation of the roles of chair and chief executive, non-executive directors, remuneration committees, nomination committees, as well as shareholder voting, and it has now become a standard expectation that firms comply. The thinking behind it is very much in terms of

[1] Investment is passive because based on a buy-and-hold strategy; ownership is active because of direct engagement with managers. See Nesbitt (1994).

standards of public conduct, for example non-executive directors have a key role because of their independence. The code has now become standard practice and it is a requirement for listing that companies provide a statement of their policies in relation to it.

From the point of view of shareholder accountability companies can be seen as public bodies and their governance discussed in the same way as that of government bodies. Corporate governance then becomes a matter of standards of behaviour, probity and accountability.[2] Firms are expected to comply with a code of conduct in order to protect shareholders and to create the information and the conditions to enable them to act as rational economic agents in their dealings with the firm. Shareholders are the group to whom the management is accountable and voting is the mechanism by which it is enforced. Shareholder democracy is similar to political democracy in that decisions that must be taken are in the nature of public goods. If an individual shareholder makes a proposal that ensures the firm is well run and profitable the benefits accrue to all shareholders in proportion to their shares, and the individual cannot gain disproportionate personal advantage. Shareholders are not in conflict with each other about the division of the profits; that is fixed by the distribution of shares and the openness surrounding a public company. At the heart of the system of corporate governance is the private provision of a public good. An analysis of shareholder incentives along these lines is presented in Leech (2003).

Traditionally the role of shareholders has been discussed in terms of control of the company, and this chapter is no exception because control involves a level of power that is easy to understand. However, many institutional shareholders do not seek control for a variety of reasons—as reported for example in Charkham and Simpson (1999). A controlling shareholding in a large company would often be larger than the rules of risk diversification would suggest, but it would also involve the investor taking a degree of responsibility for the firm which he may be ill prepared for or not want. There is also a likelihood of the investor gaining information which would compromise his share trading activities and run the risk his being accused of the criminal offence of insider trading. Most institutional shareholders are also averse to being publicly associated with a company by being seen to have control. It is therefore much more likely to be the case that shareholders seek power in the form of influence rather than control. Given the community of interests between shareholders in a company however, the question of collective control by a group of shareholders is important to determining ownership control. Voting is a matter of the shareholders deciding between management recommendations and any counter-recommendations that may be made by owners.

[2] It is in these terms that the matter is discussed in the Cadbury Report (1992).

3 Shareholder Voting Power and Corporate Control

Berle and Means (1932) were the first to study in detail the links between the concentration of share ownership, shareholder voting power and company control. They showed that in 1929 ownership of a typical large corporation in the United States had become widely dispersed among a very large group of small shareholders. The most important shareholdings of many large corporations were found to be very small indeed—in percentage terms—often <1 percent of the voting stock. Berle and Means inferred that in such cases no shareholding could be sufficiently powerful to be able to exert any real influence with management and that therefore such corporations could not meaningfully be considered to be controlled by their owners. In the default of ownership control they were assumed to be management controlled. It should be noted in passing however that their analysis was in terms of shareholdings in percentage terms and in large corporations even quite small percentage holdings represented very substantial accumulations of wealth.

Not all public companies had such dispersed ownership however. A few had a majority controlling shareholder, often the founder, or another corporation. More common was the situation where there was a large minority shareholder; in many cases such a shareholder was found from other evidence to be able to dominate through his voting power, and in effect had control. Others were controlled by minority ownership blocks through a legal device such as pyramiding or dual class shares, rather than the voting power deriving from a large minority holding.

Berle and Means made numerous detailed case studies to establish the relationship between ownership structure and control. They examined both cases where there was a stable control regime in which a corporation had obviously been controlled by the same minority shareholder or group over a long period, and also cases where there had been a proxy fight and control had either changed hands or been reaffirmed in a proxy fight. The evidence they used was a careful reading of press reports and also the kind of detailed analyses available to investors in brokers' notes and so on, so called "Street knowledge" (i.e. Wall Street knowledge). They were particularly interested in determining the point at which a minority shareholding became so small that it was no longer able to dominate voting. They identified "working control" if a minority shareholder had "sufficient stock interest to be in a position to dominate a corporation through their stock interest" and "…the ability to attract from scattered owners proxies sufficient when combined with their substantial minority interest to control a majority of the votes at the annual elections [of directors]. Conversely this means that no other stockholding is sufficiently large to act as a nucleus around which to gather a majority of the votes."

Berle and Means' careful use of case study evidence, rather than for example a simplistic statistical analysis of voting figures at annual meetings, is particularly appropriate since they were interested in studying *power*, and power by its nature is often difficult or even impossible to observe directly. For example, the existence of control by a shareholder would never be revealed to an investigator if it were never formally challenged by proxy vote or vote at company meetings. The exercise of power might be real nevertheless, with decisions being taken by the controlling shareholder and communicated to management through informal channels. It would be just as real in the circumstance, even more difficult to observe but easy to imagine, where the management understood the controlling shareholder so well that there were no need for even informal communication, perhaps because they had worked together for so long that they were of one mind. Moreover this would be well known to all the major investors in the company, even if not immediately obvious to an outsider.

The methodological question of how to study power in general has been discussed at length by Morriss (1987). He argues that intrinsically power can rarely be observed directly and that any evidence must be used in indirect ways: "there is no easy mechanical way of establishing how much power someone has and the connection between the assertion that someone has power and the evidence for it is often complex and subtle." Consequently he argues that power should best be studied using a variety of approaches within a regime of "methodological tolerance". Specifically he maintains that research into power should not be confined to the use of "hard" evidence. His most radical proposal is that researchers should be allowed to use information gained by asking other people whose opinions might be taken as authoritative evidence: by their background or practical experience they are *experts*.

Morriss suggests that there are five general approaches to gaining evidence about power that should be used in conjunction with each other: (1) Experiments; (2) Thought experiments; (3) Natural experiments; (4) Consulting experts; (5) Resource-based approaches. The approach adopted by Berle and Means fitted into this framework, and can be thought of as a combination of (3), (4) and (5). Direct experiments were impossible. Thought experiments apply to situations where the conclusion can be worked out theoretically, for example a control structure based on a legal device is obviously different from one based on a powerful minority voting block. Berle and Means' basic data were provided by natural experiments, the usual case in empirical economics.

Resource-based approaches study power in terms of the basic resources from which it derives; in the context of ownership control the relevant resources are the shares of different shareholders, in the context of the voting rules defined in the Articles of Association, and Berle and Means' attempts to define and identify controlling shareholdings comes into this category. The manner of the distribution of the shares among different shareholders determines the power of each particularly the largest shareholder. A large minority shareholder has *control* if the remaining shares are so widely distributed among a mass of small shareholders that it is very likely to be able to determine the outcome of a vote. In general the votes of the small shareholders are likely cancel each other out and give the power of decision to the large blockholder. However, where the second largest shareholder has a large weight, this power is denied to the largest shareholder and he does not have working control. Likewise if there are a number of substantial blockholders. The analysis of control depends on the complete distribution of ownership among all shareholders (and the decision rule, which for firms is almost always a simple majority). Studying this relationship formally is the primary purpose of this chapter.

Berle and Means in effect made a lot of use of Morriss' approach of consulting experts by relying on newspaper reports and "Street knowledge" to obtain independent evidence on control and related this to their shareholding data. They did not infer that a corporation was owner controlled unless they were sure they could observe it, even if indirectly, in their case studies. They reached the conclusion that a shareholding was sufficiently large to have working control through voting power if it was larger than about 20 percent, although this could vary, in many cases the figure being rather lower or higher depending on the other shareholdings. The use of this 20 percent rule to define a shareholding as controlling is commonplace in empirical work, most recently by La Porta et al. (1999).

More recent indirect evidence on the relationship between shareholder voting and control of the "consulting experts" type is in the listing rules of the London Stock Exchange (the "Yellow Book", London Stock Exchange (1993)) which uses the term controlling shareholder for one which determines the votes of 30 % or more of the shares of the company. This official definition has been drawn up by the members of the exchange in the light of their combined wisdom and experience, as practitioners who regularly back their judgement with both their own wealth and that of others, as well as their reputations. It might be supposed therefore that it has not been done lightly and might therefore be supposed to be a reflection of the opinion of experts. It is significant that it is in terms of a minority holding with working control well under the 50 % needed for legal control.

4 The Measurement of Voting Power

Since different shareholders cast different numbers of votes according to the sizes of their holdings the analysis of voting power and control is a natural application of a *weighted majority voting game*. These games are interesting because a key property of weighted voting is that the power of each player, as the ability to influence the outcome of any particular vote, does not have a simple relation to that player's voting weight. Formally it is necessary to distinguish between voting *weight*, represented by the shareholding, and voting *power*, as the ability to swing a vote, that is, the ability to *swing* a coalition of players from losing to winning by joining it.

An example illustrating this point is a company with three shareholders whose holdings are 49, 49 and 2 %. Clearly although the weights differ considerably, one of the shareholdings being very much smaller than the others, when we consider their individual power to swing the decision, they are all equal. Any two are required for a simple majority decision: the 2 % player can join with one other to swing the vote from a minority with 49 % to a majority with 51 %,[3] and each of the two 49 % players can swing the vote from 49 % to a majority with 98 %.

Counting the number of swings each player can make gives an absolute measure of power. Taking into account also the total potential number of votes which can be taken within the game, or the potential total number of swings among all the players, enables a power index to be defined for each player. Consider first all the four possible coalitions of votes which the 2 % player could join: {Ø} (the empty set), {49}, {49}, (49, 49}, the total votes being 0, 49, 49, 98 which would become 2, 51, 51 and 100. It can therefore swing two of them, the two with 49 %; it can make no difference to the decision by voting with the coalition in the other two cases. This player can therefore swing 1/2 of the decisions so its power index is 1/2. For one of the 49 % players, the coalitions are {Ø}, {2}, {49}, {2, 49} and the total numbers of votes are 0, 2, 49, 51 which become 49, 51, 98, 100. Therefore this player with 49 % weight can swing two decisions out of 4 and therefore its index is also 1/2. Therefore each of the three players has an ability to swing 1/2. It is mathematically convenient to consider all the possible voting outcomes which could occur as if they were random and equally likely since the approach treated each equally. Therefore the probability of a swing is 1/2 for each player.[4]

By contrast, as an example which illustrates the utility of the approach, consider a company with one shareholding of 30 % and 70 shareholdings of 1 %. A decision by majority vote requires 51 % support. Consider the power of the large blockholder. There are 2^{70} different possible coalitions of the small players, since each can vote either "for the motion" or "against the motion". Assuming each small player votes each way with equal probability independently of the others, the total number of votes cast by them "for the motion"—call this Y—is distributed with a binomial distribution, with parameters (in the usual notation) $n = 70$ and $p = 0.5$, or in the usual shorthand, $Y \sim B(70, 0.5)$. The swing probability of the large player is then found using this distribution, as the probability that the large player can swing the vote, which occurs when Y is at least 21 and <51. This is the binomial probability, $P(21 \leq Y \leq 50) = 0.999370$. Therefore the 30 % player is very powerful, in that his swing probability is very close to unity indeed, but it is necessary to check the powers of the small players also to establish relative power.

So consider a player with 1 % of the votes. A swing occurs when that player is able to change a losing coalition into a winning one, which means changing one with 50 % of the votes into a 51 % majority. In this case it is necessary to consider the total votes of 69 small players as random and also to treat the votes of the largest player as being random. The total number of votes cast by the small

[3] The decision rule requires a 51 % majority here because the examples involve discrete data. The analysis of the real data later in the chapter will use a 50 % rule.

[4] There are three players each with a power index of 1/2. In the literature on power indices it is frequently assumed that the total power of decisions is divided among the players so that the indices represent shares of power and sum to one. In this example if such a *normalised* index were used each player would have an index of 1/3. I do not adopt this approach for reasons discussed below, following Coleman (1971).

players, say U, has the binomial distribution, $U \sim B(69, 0.5)$. To find the swing probability of a small player with 1 % of the votes it is necessary to allow for the possible behaviour of the large player as well as the other 69 small players. There are two equally probable cases: (1) where the large player votes "for", so therefore for a swing $30 + U = 50$, and so we must have $U = 20$; (2) where the large player votes "against" so therefore $U = 50$ for a swing. The swing probability for the small player is then $0.5P(U = 20) + 0.5P(U = 50) = 0.000137$.

It is clear from this example that the player with 30 % is effectively totally dominant and has very close to complete control, while the small players individually are virtually powerless. This property of weighted voting to assign very great power to a block of votes faced by a very dispersed distribution among a large number of other players explains why shareholder power is so important to the system of corporate governance even in countries with dispersed ownership like the UK. Dispersed ownership in itself does not necessarily imply dispersed power.

The idea of a power index as a general measure of voting power originated in the classic paper by Shapley and Shubik (1954 and 1988).[5] The Shapley-Shubik index proposed there was an application of the Shapley value (Shapley (1953 and 1988)) as a method of evaluating the worth to each player of participating in a game. The central idea of the Shapley value was *bargaining* among the players over the spoils of a decision. This bargaining approach to thinking about voting in a collectivity was however severely criticised by Coleman (1971) who argued that the consequences of a collective decision taken by majority voting could not usually be thought of in this way. A decision about an action that the collectivity could take would have consequences for the members that could only be understood in the wider context, and could not be conceived of as sharing the spoils. An example would be a decision to replace the top management in a public company: if performance subsequently improved entitlement to the additional profits would normally be distributed among all shareholders in proportion to their shareholdings and not according to their individual voting powers.

The alternative approach therefore is one in which the outcomes are in the nature of public goods; voting is a matter of political democracy and the power index is a measure of general voting power and not a value. Coleman advocated an approach in which the voting body is analysed in terms both of the powers of voting members[6] and also the power of the body itself to act. Banzhaf (1965)

---

[5] Shareholder voting was always suggested as an application of these ideas, right from the earliest days, see Shapley (1961).

[6] He proposed separate measures of the power to *initiate* action and the power to *prevent* action but this distinction only matters for bodies which employ a supermajority. When the decision rule requires only a simple majority for a decision these two indices are equivalent. For this reason Coleman's approach has tended to be dismissed as equivalent to that of Banzhaf and there have been few if any applications of it. Coleman argued forcefully against the idea of a power distribution in which the total power of decision making is shared out, which is a central idea in the Shapley-Shubik index. The swing probabilities used in the current chapter can be thought of as Coleman's powers to initiate action but I also use *normalised* power indices, or Banzhaf indices, to measure, not shares of power, but relative powers of different players.

proposed an index of power in weighed voting situations based on a different coalition model from that of Shapley and Shubik—the model that I have described above. Both these indices are often referred to in the literature as the classical power indices and both have been widely applied with sometimes similar but often widely different results. This has led to a problem of choice of index and, in the absence of independent evidence on the powers of players in the real-world weighted voting games to which they have been applied,[7] to something of an impasse in the development of the field. This has prompted considerable theoretical work on the comparative properties of the indices, to the proposal of new indices, and also to the rejection of the power indices approach entirely. Nevertheless the method promises to have utility in the analysis of power in general voting systems and in the design of constitutions.

Accounts of the measurement of power and of the different indices and the theoretical debates on their comparative properties are given in Lucas (1983), (Straffin and Philip 1994) and Felsenthal and Machover (1998). An empirical comparison of the two classical indices which clearly suggests the inadequacy of the Shapley-Shubik index is reported in Leech (2002a) and on the basis of that analysis this chapter will confine itself to the use of the Banzhaf (non-normalised version) or Coleman index, which will be referred to below simply as a power index. The details of the calculation of the indices are omitted. They are given in Leech (2001).[8]

5 The Applicability of Power Indices to Shareholder Voting

The approach to the measurement of power just described treats the firm as a public body regulated by high standards of corporate governance including the legal protection of shareholder rights, rather than simply a source of profits to be split among the owners by bargaining based on power, a model perhaps more appropriate to private companies. The question arises as to whether the measure of power used is appropriate in this context given its assumptions. The power index is a measure of abstract power and has no regard for preferences or the issues about which voting takes place. This is obviously something that has to be qualified since it will not apply in all cases. It can not be applied to issues on which all shareholders are unanimous, such as a policy which makes them unambiguously happier

[7] For example, the United Nations, the US Presidential Electoral College, and the European Union Council.

[8] In a previous chapter (Cubbin and Leech (1983 and 1999)), John Cubbin and I proposed a measure of the voting power of the largest shareholding block which we called the *degree of control*. The degree of control was defined as the probability that the largest block could be on the winning side in a vote, assuming the same voting model as the power index. There is a simple relation between it (denoted by DC) and the power index for the largest shareholder, $PI_1 = 2DC - 1$.

or one that reduces the value of the firm with no offsetting benefits. Nor can this model be used to make statements about control involving a powerful minority shareholder being able to expropriate the majority by appropriating the private benefits of control to himself.

The approach adopted in this chapter is one where the firm is regarded as a democratic body that has to make strategic decisions in situations of fundamental uncertainty where the potential for making mistakes is enormous. There are many situations where this occurs. For example, a retail company may have enjoyed considerable success in expanding its sales of a new brand and have developed a chain of very profitable shops. The chief executive may wish to build on this success by an ambitious policy of expansion on a much larger scale and proposes the purchase of a large store, much larger than any in the chain, in the centre of every major city in the country. Extrapolating past performance, the proposal would seem to be profitable, but the quantum change in scale involved raises the question of whether the formula that has been successful in the past would still continue to be so. Another example would be where a successful business expands abroad; there are many examples of British and German companies that have lost out by attempting to expand into the United States.

Other examples occur where changes in the external trading environment take place which necessitate a fundamental strategic reappraisal. An example would be a successful clothing retailer which develops its own credit card primarily for use in its stores; demand for clothes falls as the market for clothing changes with changing consumer tastes leaving the company with a profitable financial services division but no longer a profitable clothing seller. Shareholders will inevitably have to decide between two incommensurable strategies: on the one hand, changing the fundamental nature of the business from primarily selling clothes to financial services, and on the other, a new management plan confidently proposed which will guarantee to restore former glory. A common case is where the board of directors is split, the management on one side and the non-executive directors on the other, the shareholders having to resolve the issue.

Another example that occurred recently in the UK is where there two rival bids to take over a company, which may differ in the bid price but are also different in the method of financing. Both bids are in terms of a mixture of cash and shares but the higher bid has a higher share element and there is uncertainty about what the share value will be. In such a case the model of shareholder voting applies since there is no objective reason to vote either way in the absence of information. Another case where the model might apply is where the chief executive wishes to be paid a large rise on promises of future success; shareholders must decide this on the basis of unknowable future performance. Where there is always this kind of uncertainty is in the appointment of directors and especially the chief executive; there may be two candidates with similar track records and there may be strong reasons for appointing each, but there may turn out to be large differences in competence in the future were either to be appointed.

In all such cases, the voting model used to measure shareholder power is a reasonable approximation and also the voting power of large shareholders is important in determining the outcome. Shareholders usually have to decide whether to accept management proposals to enhance shareholder wealth which also benefit management. Often the benefit obtained by management is in the short run and that by shareholders over a much longer term. In the absence of substantial share ownership by management, which is a reasonable assumption since directors holdings are no longer significant in the great majority of companies in the UK, there is little difference of interest among shareholders, and therefore shareholders are not likely to be committed to any particular side in the vote.

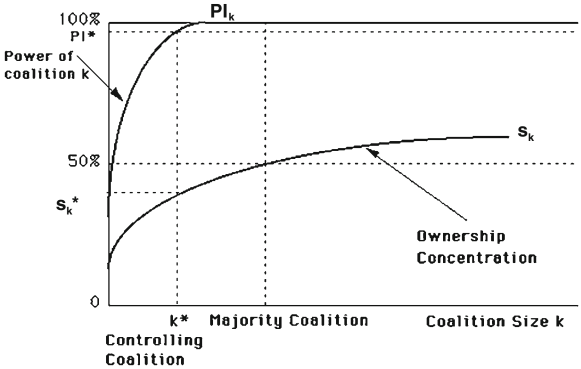
6 A Model of Ownership Control

In previous work Leech (1987) I proposed a model of minority ownership control based on the formal voting power of the largest block of shares as measured by a power index or the degree of control. A company is classified as owner-controlled if the power index for the largest shareholder or group of shareholders exceeds some very high level and no other has any appreciable voting power. The essential advantage of this approach over the conventional "fixed rules" approach to determining control used by many authors[9] is that the power of a large ownership block depends not only on its percentage of the voting equity but also on the dispersion of the other shareholdings. The fixed rule infers control only from the size of the largest block. Thus, for example, a shareholder with 20 % of the shares could be regarded as controlling in some cases but not in others on the basis of power indices, while it would always be deemed to be controlling if a fixed 20 % rule were used.

Figure 1 shows the model of minority voting control described in Leech (1987). The horizontal axis shows the number of members of the potential controlling group, starting with the largest and adding successively smaller holdings. Let the block consisting of the k largest shareholdings comprise s_k shares and its corresponding voting power be measured by its power index, PI_κ; both functions are shown on the vertical axis. A typical concentrated ownership structure is shown with the ownership-concentration function s_k and the power-index function PI_k. The block has majority control when it has k' members, such that $s_{k'} = 0.5$ and therefore $PI_{k'} = 1$. The block is assumed to have minority control when its power index is very close to 1. In the diagram this is represented as being when the block size is k* members and its voting power is PI*. The threshold PI* is chosen appropriately. This model is the basis of the empirical approach reported in the

[9] See Short (1994) for a survey. La Porta et. al. (1999) have recently used a fixed rule based on 20 %.

Fig. 1 A model of "minority control"



next section.[10] Since the model is being used here to examine properties of the distribution of ownership, and the blocks are theoretical rather than actual, in the results section below they are referred to as "controlling" in quotes.

7 The Data Set: Large Voting Shareholdings in a Sample of Large UK Companies

The data set is based on the sample collected by Leech and Leahy (1991). It consists of those companies, 444 in number, where there was no majority shareholder. All were listed on the London Stock Exchange in the mid-eighties and included about a third of the Times 1000 as well as some smaller companies and some financial companies. They comprise neither a representative sample nor a random sample since they were chosen on the sole basis of the availability of detailed ownership data to give the voting weights. The source was a commercial information service, which existed for a short time, called "Who Owns What on the London Stock Exchange", to which one could make an annual subscription and receive periodic printouts showing details of all shareholdings greater in size than 0.25 % of the total of each class of equity.[11]

For most companies there was only one class of voting share but in the small number of cases where there were two, they were combined into one distribution taking into account any differences in voting weights and voting rules. The source of the information provided by "Who Owns What on the London Stock Exchange", was company share registers maintained under the Companies Act legislation, made publicly accessible in some form at Companies House and

[10] There is a potential identification problem here since the model can be used to determine control endogenously by choosing the shape of the curve s_k. Therefore we might expect observed ownership structures of actual firms to reflect this.

[11] The Warwick University Library took out a 1-year subscription to it at my suggestion.

Table 1 The sample: The largest holding versus the second largest

		w_1						
		<5 %	5–10 %$	10–20 %	20–30 %	30–40 %	40–50 %	Total
		41	144	125	85	30	19	444
w_2	<5 %	41	46	15	12	2	2	118
	5–10 %		98	73	26	10	9	216
	10–20 %			37	38	11	5	91
	20–30 %				9	4	2	15
	30–40 %					3	1	4
	40–50 %						0	0

periodically searched by the service. This provides a much richer data set than the declarable stakes of 3 % or more that companies are obliged to publish in their Annual Reports (the basis of many studies of ownership and control) but requires much more processing before it is useable. Many of the holdings were in the names of nominee companies but wherever possible these were reassigned to their beneficiaries using a directory of nominees provided with the subscription to identify them. Holdings in the same firm by different members of the founding family, and other interest groups closely associated with the company, were amalgamated into a single block using surnames and other information. The data used therefore can be assumed to be reasonably close to beneficial holdings taking into account voting alliances.[12]

The data collected were based on searches of company registers made in 1985 and 1986. The number of large shareholdings observed (after amalgamation by Leech and Leahy) varies in the sample between a minimum of 12 and a maximum of 56, with a median of 27. The proportions of voting equity these represent vary between 19 and 99 %, the median being 66 %. The dataset is therefore both detailed and fairly comprehensive.

The data are summarised in Table 1. The table shows the distribution of the size of the largest shareholding, w_1, and also the joint distribution of w_1 with the second-largest holding, w_2, in order to indicate the variation in patterns of ownership concentration between firms in the sample. Some 49 companies have relatively concentrated voting structures with w_1 greater than 30 %, but in the great majority of cases w_1 is less than 30 %. There is also a wide range of variation in the size of w_2 given w_1. For example in the group of 85 companies where w_1 is between 20 and 30 %, w_2 is less than 10 % in 38 cases, between 10 and 20 % in a further 38 cases and greater than 20 % in 9 cases.

[12] The source and method of construction of the data set are described in Leech and Leahy (1991). There might remain a slight underestimation of the true concentration of ownership to the extent this information was incomplete.

8 The Problem of Incomplete Data

The data collected on the distribution of share ownership is necessarily incomplete because large public companies typically have many thousands of shareholders and it would be prohibitively costly to collect them all. In any case, in practice almost all of these are too small to have any real individual voting power and little would be gained by going to the trouble of collecting the data. On the other hand, however, they have a formal role to play in the voting games being assumed in this chapter and therefore it is necessary to deal with them appropriately.

The solution to this incompleteness problem adopted here is to analyse two modified games for which the data we do have would be appropriate. Two sets of indices are calculated, assuming two different games where the unobserved players conform to two extremes of "concentrated" and "dispersed" ownership. These are both arithmetically consistent with the observed data. The "concentrated" case takes the extreme that the unobserved weights are all equal to the threshold for observation, 0.25 % (however, strictly slightly smaller) and the number of players is finite if large. The "dispersed" case assumes an "oceanic game" where the unobserved small holdings are taken to the limit where each of them is individually infinitesimally small and they are infinite in number.

Thus, for any company, say k shareholdings are observed out of a total of n. The shareholdings or voting weights are represented by the notation w_1, w_2, w_3, etc. in decreasing order of size, starting with the largest, and the smallest is w_k which is normally equal to 0.0025. There is no information about the remaining n-k holdings except that they are all no larger than w_k. Nor is it necessary to know n; although the total number of shareholders could be collected from share registers, it would add very little to the analysis. The two limiting cases are referred to respectively as limiting case C (Concentrated) and limiting case D (Dispersed). For limiting case C it is necessary to adopt a value for n in the finite game. If w_k is the smallest weight observed in the data, then all the non-observed weights are no greater than w_k. The most concentrated pattern of ownership occurs when they are all equal to w_k. Then the corresponding value of n, call it n', is: $n' =$ integer part $((1 - s_k)/w_k) + k + 1$ and we let $w_i = w_k$ for all $= k + 1,\ldots, n'-1$ and $w_{n'} = 1 - s_k - (n' - k - 1)w_k$. Obviously $w_k = 0.0025$.

These two cases are analysed separately as different games, case C as a finite game using the algorithm described in Leech (2001) to calculate the indices and case D as an "oceanic" game. Power indices for oceanic games have been thoroughly studied and there is a good literature on them. The approach adopted here follows that of Dubey and Shapley (1979), who showed that the power indices for an oceanic game with k major players with combined weight of s_k and a majority requirement or quota of q are the same as for a finite game consisting only of the k major players and a modified quota of $q - (1 - s_k)/2$. These can be calculated using the algorithm of Leech (2001).[13]

9 Power Indices for Illustrative Companies

Table 2 presents power indices for large shareholdings in some illustrative companies. The firms have been selected to span the range of variation in the first two shareholdings within the sample. Plessey has the most dispersed ownership with a largest shareholding of under 2 % and Associated Newspapers is one of several which are just short of having majority control. Two firms have been selected in each range of values for w_1: 10–20 %, 20–30 %, 30–40 %, 40–50 %. In each range the two companies are those with relatively large and small values for w_2. The results for these firms might then be taken as illustrative of the effects of ownership concentration in terms both of the size of the largest holding and the relative dispersion of the other holdings as reflected in the second largest. Results are shown for representative shareholders numbered 1, 2, 3, 5, 10 and 20.

The values of the power indices in Table 2 are sensitive to differences in ownership structure and vary considerably. They appear to conform to commonly held a priori notions of the power of shareholding blocks of a given size in relation to others. Where ownership is widely dispersed as in the case of Plessey, power is also widely dispersed. Where it is highly concentrated, as in Ropner or Steel Brothers, with a shareholding over 40 %, giving control, the index reflects this. In other cases where ownership is less concentrated, there is considerable variety of results associated with differences in ownership structure.

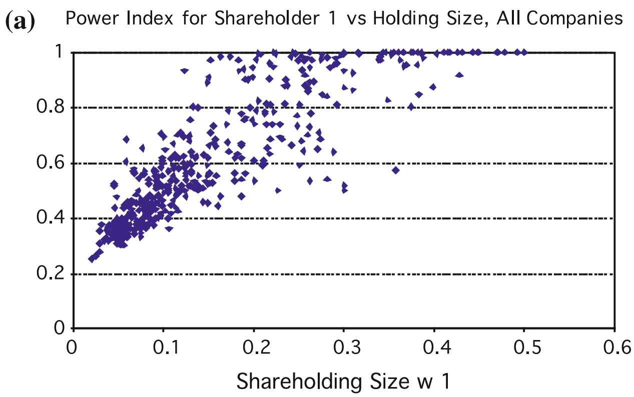
A comparison of Sun Life and Liberty, for example, shows the sensitivity of the power of the largest shareholder to the size of the second largest shareholding. The 22 % largest shareholding in Sun Life has a power index over 99 % suggesting that it can be regarded as a controlling holding and reflecting the relatively high dispersion of ownership of the other 78 % of shares. In the case of Liberty, however, both the largest two holdings are above 22 % which must mean that the largest shareholder is not much more powerful than the second-largest and this result is obtained; both have an index of about 0.5 and, in this case, the third shareholder has enhanced power as a result. A similar finding emerges for companies with a shareholding of between 30 and 40 %. A 31 % shareholding has a power index over 99 % in Securicor where there are no other large owners. On the other hand a similar-sized stake in Bulgin has an index of only 86 % because of the presence of a large second shareholder with 22 % of the votes.

[13] Typically the finite games assumed for case C have upwards of 300 players and require an algorithm which can cope with such large games. As regards the oceanic games in case D, the results of Dubey and Shapley are subject to conditions on q to ensure existence, but in this case $q = 0.5$ and the conditions are always met.

Table 2 Power indices for top shareholders, illustrative companies

Company	Shareholder:	1	2	3	5	10	20
Plessey	Weight	0.019	0.015	0.013	0.011	0.009	0.004
	Index (C)	0.254	0.192	0.165	0.134	0.112	0.052
	Index (D)	0.361	0.268	0.230	0.185	0.154	0.071
United spring & steel	Weight	0.123	0.109	0.098	0.037	0.014	0.005
	Index (C)	0.502	0.433	0.391	0.117	0.046	0.016
	Index (D)	0.508	0.440	0.400	0.113	0.045	0.016
Suter	Weight	0.128	0.065	0.053	0.031	0.017	0.009
	Index (C)	0.692	0.246	0.209	0.120	0.068	0.034
	Index (D)	0.707	0.244	0.210	0.121	0.068	0.034
Ranks Hovis McDougall	Weight	0.149	0.037	0.035	0.022	0.014	0.008
	Index (C)	0.912	0.070	0.068	0.047	0.031	0.017
	Index (D)	0.940	0.053	0.052	0.038	0.025	0.014
International signal & control	Weight	0.163	0.032	0.018	0.016	0.011	0.004
	Index (C)	0.984	0.015	0.011	0.010	0.007	0.003
	Index (D)	0.998	0.002	0.002	0.002	0.001	0.001
Sun life	Weight	0.222	0.035	0.019	0.013	0.009	0.005
	Index (C)	0.9996	0.0004	0.0003	0.0003	0.0002	0.0001
	Index (D)	1.000	0.000	0.000	0.000	0.000	0.000
Liberty	Weight	0.2263	0.2257	0.089	0.050	0.018	
	Index (C)	0.5013	0.4982	0.278	0.132	0.047	
	Index (D)	0.5014	0.4983	0.280	0.133	0.047	
Securicor	Weight	0.316	0.073	0.053	0.029	0.016	0.008
	Index (C)	0.997	0.003	0.003	0.003	0.002	0.001
	Index (D)	0.998	0.002	0.002	0.002	0.001	0.001
Bulgin	Weight	0.310	0.222	0.045	0.028	0.009	0.003
	Index (C)	0.862	0.138	0.122	0.079	0.025	0.007
	Index (D)	0.874	0.126	0.120	0.082	0.025	0.007
Ropner	Weight	0.410	0.060	0.050	0.020	0.012	0.003
	Index (C)	1.000	0.000	0.000	0.000	0.000	0.000
	Index (D)	1.000	0.000	0.000	0.000	0.000	0.000
Steel brothers	Weight	0.425	0.213	0.038	0.030	0.007	0.003
	Index (C)	0.9996	0.0004	0.0004	0.0004	0.0002	0.0001
	Index (D)	0.9999	0.0001	0.0001	0.0001	0.0000	0.0000
Associated newspapers	Weight	0.4995	0.026	0.021	0.021	0.013	0.006
	Index (C)	1.0000	0.000	0.000	0.000	0.000	0.000
	Index (D)	1.0000	0.000	0.000	0.000	0.000	0.000

These results are plausible in that they are in broad agreement with both the results of Berle and Means (1932) and more recent conventional ideas about the power of shareholder blocks and minority ownership control. It has been possible to find cases where the power index for a voting block greater than 20 % is extremely close to 100 %.



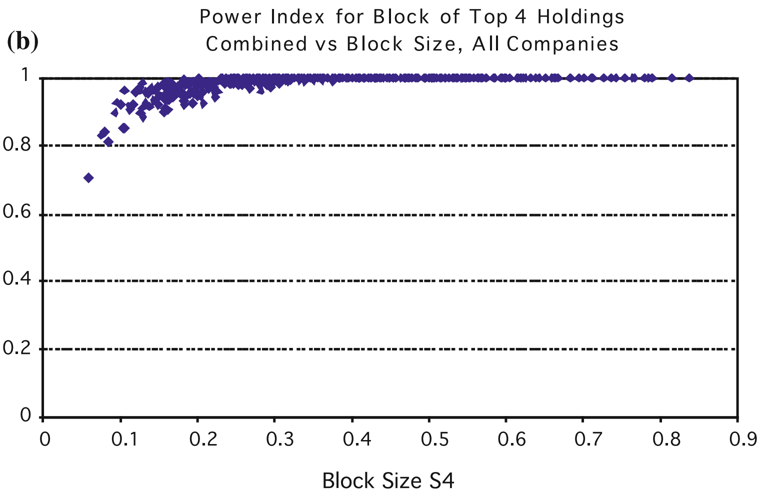


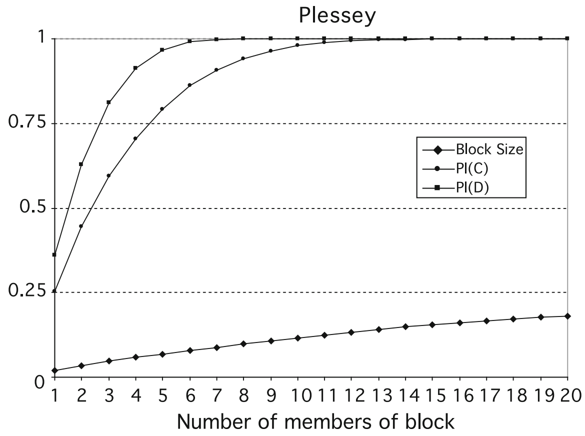
Fig. 2 **a** Power index for shareholder 1 versus holding size, all companies. **b** Power index for block of top 4 holdings combined versus block size, all companies

10 The Complete Sample

Results for the full sample are shown in Fig. 2. Figure 2a shows the respective power indices for the largest shareholding, PI_1, against its size w_1; Fig. 2b shows the equivalent plots after the largest 4 shareholdings have been combined into a single block, of size s_4. Only the results for case C have been presented since the oceanic indices are very close. These plots are useful for giving an insight into the respective behaviour of the power indices in the population as a whole and their potential as a basis for identifying minority control.

There is considerable variation reflecting differences in ownership structure. Concentration in terms of the size of the largest shareholding has very little effect

Fig. 3 The power of a bloc of large shareholders



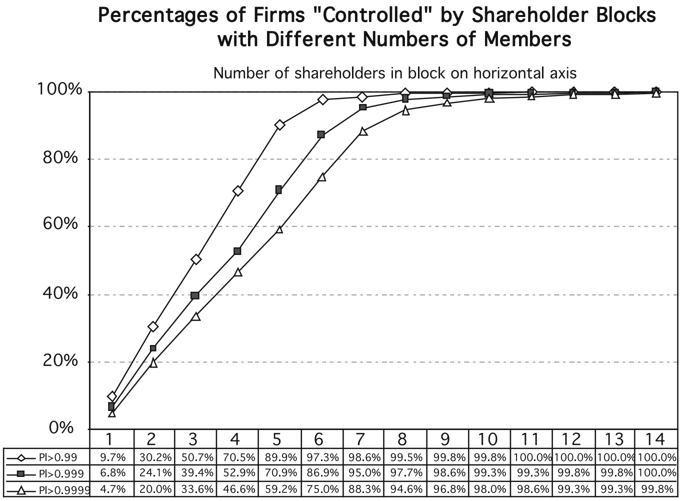
up to over 15 % but after that power varies widely. These results suggest that shareholdings between 20 and 30 % can be said to have voting control in many cases but not in many others. Voting control is possible on the basis of a holding below 20 % but such cases are not common. Most (but not all) holdings greater than 35 % have a power index equal to or almost equal to 1. The variation suggests that this index may be useful as a guide to control on the basis of individual shareholding data.

Figure 2b shows that combining the top four shareholdings into one voting block is very powerful indeed in most cases. In some companies such blocks would be majority shareholders but it is interesting that the result does not depend on this. Intuitively combining top shareholdings has a double effect in both increasing concentration via the size of the block and reducing the dispersion of the remainder; these reinforce one another in concentrating power.

11 Potential Controlling Blocks

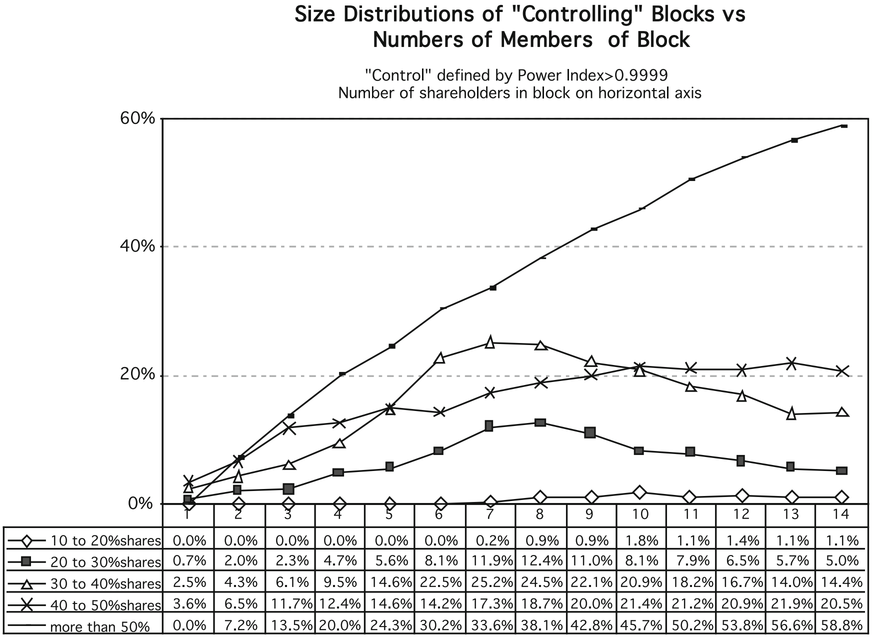
Figure 3 examines the model of ownership control by a block of large shareholders presented above in the light of data, by graphing the power of blocks of different sizes. Results are shown for illustrative companies in which the power indices have been calculated for each assumed block of shares, of size s_k, for $k = 1$ to 20,[14] the ownership concentration curve and the power curve. Plots are given for two companies, Plessey, which has the most dispersed ownership structure, and Birmid Qualcast, only slightly more concentrated. Each plot shows the number of members of the group, k, on the horizontal axis and s_k, the size of the block, and the associated power index on the vertical axis. The plots show the same general

[14] 16 for Liberty.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14
PI>0.99	9.7%	30.2%	50.7%	70.5%	89.9%	97.3%	98.6%	99.5%	99.8%	99.8%	100.0%	100.0%	100.0%	100.0%
PI>0.999	6.8%	24.1%	39.4%	52.9%	70.9%	86.9%	95.0%	97.7%	98.6%	99.3%	99.3%	99.8%	99.8%	100.0%
PI>0.9999	4.7%	20.0%	33.6%	46.6%	59.2%	75.0%	88.3%	94.6%	96.8%	98.0%	98.6%	99.3%	99.3%	99.8%

Fig. 4 Potential controlling blocs. **a** Percentages of firms "Controlled" by shareholder blocks with different numbers of members. **b** Number of shareholders in block on horizontal axis



	1	2	3	4	5	6	7	8	9	10	11	12	13	14
10 to 20%shares	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.9%	0.9%	1.8%	1.1%	1.4%	1.1%	1.1%
20 to 30%shares	0.7%	2.0%	2.3%	4.7%	5.6%	8.1%	11.9%	12.4%	11.0%	8.1%	7.9%	6.5%	5.7%	5.0%
30 to 40%shares	2.5%	4.3%	6.1%	9.5%	14.6%	22.5%	25.2%	24.5%	22.1%	20.9%	18.2%	16.7%	14.0%	14.4%
40 to 50%shares	3.6%	6.5%	11.7%	12.4%	14.6%	14.2%	17.3%	18.7%	20.0%	21.4%	21.2%	20.9%	21.9%	20.5%
more than 50%	0.0%	7.2%	13.5%	20.0%	24.3%	30.2%	33.6%	38.1%	42.8%	45.7%	50.2%	53.8%	56.6%	58.8%

Fig. 5 Sizes of potential controlling blocs. **a** Size distributions of "Controlling" blocks versus numbers of members of block. **b** "Control" defined by power index >0.9999 number of shareholders in block on horizontal axis

pattern for both companies, consistent with the theoretical Fig. 1, and the inference can be drawn that for the great majority of companies a block comprising a small number of top shareholders would effectively have control.

Figure 4 investigates this effect by calculating the proportion of the sample which would satisfy the definition of control by blocks of different numbers of shareholders on different definitions of control, PI* = 0.99, 0.999 and 0.9999. It shows that it is pervasive and that the power of a shareholder block comprising, say, the top six holdings would be very considerable indeed in most companies. On the control criterion of PI* = 0.9999, the model would deem over 75 % of the companies in the sample to be owner controlled. Virtually the whole sample would be owner-controlled by the top ten shareholders combined.

Figure 5 shows the size distribution of these "controlling" blocks in terms of the concentration of ownership they represent using the PI* = 0.9999 criterion. It shows that the effect reported in the previous two paragraphs does not depend on the blocks having a voting majority. For example, continuing with blocks comprising just the top six shareholders (which are deemed to control 75 % of the sample companies), in only 30 % of companies is the block a majority, and in 22 % of cases it is between 30 and 40 % of the equity. On the other hand, it represents between 20 and 30 % of the equity in only 8.1 % of cases.

12 Conclusions

This chapter has looked at the voting power of large shareholders in the widely dispersed ownership observed on the stock market of the United Kingdom. It has adopted a methodology due to Berle and Means (1932) supplemented by the technique of power indices for measuring power derived from game theory. The empirical findings are consistent with earlier work and also institutional practice.

The results show that a significant minority shareholder can be very powerful, almost as powerful as a majority shareholder, if the dispersion of the rest of the holdings is sufficient. In most companies a 20 % shareholding can have working control, but in other companies the figure is greater and in some less. In almost all companies if the top shareholders formed a voting block this would be extremely powerful. In almost all companies the top six shareholders could form a controlling voting block, whether or not it contained a majority of the shares.

The approach has treated the company as a quasi-political body in which shareholders are voters choosing public goods, a reasonable way of looking at a public company where there are good standards of corporate governance. It ignores completely the question of incentives. A better model might be one which recognises that shareholders are of two types: those with substantial stakes who have strong private incentives to take part in collective action and those whose stakes are so small that their best strategy is to abstain. This requires a model of incentives and is the subject of future work. However such a model of voting power would be likely to show that relatively small holdings are in fact very powerful within the reduced group of active shareholders that would be identified. The approach adopted here, where all shareholders are taken into account regardless of size, biases the analysis away from finding considerable shareholder power and therefore makes the results more significant.

Acknowledgments I am grateful to many people for comments that have led to improvements in the chapter but especially the editor, and two anonymous referees. Responsibility for the final version is mine alone however.

Appendix: 1 Proof that the power curve is concave

To show this, consider a block consisting of the largest k shareholders with combined shareholding $s_k = \sum_{i=1}^{k} w_i$. The power index for the block is PI_k defined as the swing probability for the coalition k in the voting model in which the votes of shareholders k + 1, k + 2,…,n are treated randomly as defined. Let x_i be the number of votes cast by shareholder i. Then x_i has the probability distribution, $Pr(x_i = w_i) = Pr[x_i = 0] = 1/2$, independently for all i. Define the random variable $Y = \sum_{i=k+2}^{n} x_i$. The swing probability PI_k can be written:

$$PI_k = 0.5\ P_r[0.5 - s_k < Y < 0.5] + 0.5\ P_r[0.5 - s_k < Y + w_{k+1} < 0.5].$$

Denoting the cumulative probability distribution function for Y by the function P(Y), this can be written as,

$$PI_k = [P(0.5) - P(0.5 - s_k) + P(0.5 - w_{k+1}) - P(0.5 - s_{k+1})]/2.$$

Now consider the index for coalition k + 1 of size s_{k+1}: $PI_{k+1} = P(0.5) - P(0.5 - s_{k+1})$ Therefore the change in the index is:

$$PI_{k+1} - PI_k = [P(0.5) - P(0.5 - w_{k+1}) + P(0.5 - s_k) - P(0.5 - s_{k+1})]/2.$$

This expression is always non-negative if $w_{k+1} \geq 0$. It is decreasing as $w_{k+1} \to 0$, since $P(0.5) \to P(0.5 - w_{k+1})$ and $P(0.5 - s_k) \to P(0.5 - s_{k+1})$. Therefore the power curve is concave increasing as drawn in Fig. 1.

References

Banzhaf, J. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review, 19*, 317–343.

Berle, A. A., & Means, G. C. (1932). *The modern corporation and private property*. New York: Harcourt, Brace and World, Inc. Revised Edition, 1967.

Cadbury Report. (1992). *The financial aspects of corporate governance*. London: Gee and Co Ltd.

Charkham, J., & Simpson, A. (1999). Fair shares, Oxford: Oxford University Press.

Coleman, J. S. (1971). "Control of Collectivities and the Power of a Collectivity to Act," In Lieberman (ed.), *Social choice, gorden and breach* (pp 277–287); reprinted in J. S. Coleman (ed.). (1986). *Individual interests and collective actions*, Cambridge: Cambridge University Press.

Cubbin, J. & Leech, D. (1983 and 1999), "The effect of shareholding dispersion on the degree of control in british companies: Theory and measurement". *Economic Journal, 93*(June), 351–369; reprinted in K. Keasey, S. Thompson & M. Wright (Eds.). (1999). Corporate governance. *Edward Elgar: Critical Writings in Economics, 2*, 61–80.

Dubey, P., & Shapley, L. S. (1979). Mathematical properties of the Banzhaf power index. *Mathematics of Operations Research, 4*, 99–131.

Felsenthal, D. S. & Machover M. (1998). The measurement of voting power, Edward Elgar.

Gambarelli, G. (1994). Power indices for political and financial decision making: a review. *Ann Oper Res, 51*, 165–173.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. W. (1999). Corporate ownership around the world. *Journal of Finance, 32*, 1131–1150.

Leech, D. (1987). Ownership concentration and the theory of the firm: A simple-game-theoretic approach. *Journal of Industrial Economics, 35*, 225–240.

Leech, D. (1988). The relationship between shareholding concentration and shareholder voting power in british companies: A study of the application of power indices for simple games. *Management Science, 34*, 509–527.

Leech, D. (2001). "Computing power indices for large weighted voting games," Warwick economic research paper number 579, revised July 2001.

Leech, D. (2002). An empirical comparison of the performance of classical power indices. *Political Studies, 50*, 1–22.

Leech, D. (2003). "Incentives to Corporate Governance Activism". In M. Waterson (Ed.), *Competition, monopoly and corporate governance, essays in honour of Keith cowling, Edward Elgar,* chapter 10, pp. 206–227.

Leech, D., & Leahy, J. (1991). Ownership structure, control type classifications and the performance of large British companies. *Economic Journal, 101*, 1418–1437.

London Stock Exchange. (1993). *The listing rules, (the yellow book)*. London: The Stock Exchange.

Lucas, W. F. (1983). Measuring power in weighted voting systems. In S. Brams, W. Lucas, & P. Straffin (Eds.), *Political and related models*. London: Springer.

Morriss, P. (1987). *Power: A philosophical analysis*. Manchester, UK: Manchester University Press.

Nesbitt, S. L. (1994). Long-term rewards from shareholder activism: A study of the CalPers effect. *Journal of Applied Corporate Finance, 19*, 75–80.

OECD (1998, April). Corporate governance, A report to the OECD by the business sector advisory group on corporate governance.

Pohjola, M. (1988). Concentration of shareholder voting power in Finnish industrial companies. *Scandinavian Journal of Economics, 90*, 245–253.

Rapaport, A. (1998). *Decision theory and decision behaviour* (2nd ed.). London: Macmillan.

Rydqvist, K. (1986). *The pricing of shares with different voting power and the theory of Oceanic Games*. Stockholm: Economic Research Institute, Stockholm School of Economics.

Shapley, L. S. (1953 and 1988) "A value for n-Person games". In H. W. Kuhn & A. W. Tucker (eds.), *Contributions to the theory of games, annals of mathematics studies, 28* (vol. II, pp. 307–317). Princeton, New Jersey: Princeton University Press; reprinted in A. E. Roth. (1988). *The shapley value*. Cambridge: Cambridge University Press.

Shapley, L. S. (1961). Values of large games III: A corporation with two large stockholders, RM-2650, Th Rand Corporation, Santa Monica, California; included in J. S. Milnor & L.

S. Shapley (1978). “Values of large games II: Oceanic games”. *Mathematics of Operations Research, 3*, 290–307.
Shapley, L. S. & Shubik M. (1954 and 1988). “A method for evaluating the distribution of power in a committee system.” *American Political Science Review, 48*, 787–792; reprinted in Roth, A. E. (1988). *The shapley value*. Cambridge: Cambridge University Press.
Shleifer, A., & Vishny, R. W. (1997). A survey of corporate governance. *Journal of Finance, 52*, 737–783.
Short, H. (1994). Ownership, control, financial structure and the performance of firms. *Journal of Economic Surveys, 8*, 203–249.
Straffin, P. D. (1994). “Power and Stability in Politics.” In R. J. Aumann & S. J. Hart (ed.), *Handbook of game theory* (Vol. 2, pp. 1128–1151). Amsterdam: Elsevier.

Part V
Voting Power in the European Union

Calculus of Consent in the EU Council of Ministers

Hannu Nurmi, Tommi Meskanen and Antti Pajala

1 Introduction

In their classic treatise *The Calculus of Consent: Logical Foundations of Constitutional Democracy* Buchanan and Tullock analyze the fundamental questions of the emergence of political systems in economic terms. The basic tenets of the book are, firstly, that "if no collective action is required, there will be no need for a political constitution" and, secondly, that "the individual will find it profitable to explore the possibility of organizing an activity collectively when he expects that he may increase his utility" (Buchanan and Tullock 1962, 43). The utility may ensue from collective action in two ways:

- the collective action may help the individual to avoid some external costs that result from the actions of others, or
- the collective action may have beneficial external effects upon the individual.

The voter calculus can be described by the following Fig. 1.

This article has benefited from comments of the editors and anonymous referees of HOMO OECONOMICUS, Matti Wiberg and the participants of the Third Biannual Meeting on Group Decision Making and Power Indices held in Józefów, Poland, July 20–27, 1996. An earlier version of this article was published in HOMO OECONOMICUS 17(1/2)

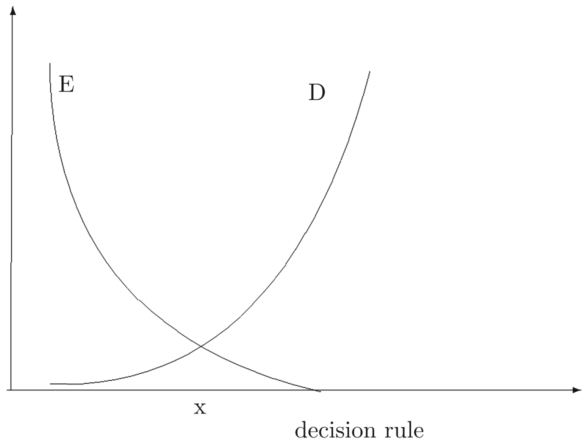
H. Nurmi (✉)
Department of Political Science and Contemporary History and Public Choice Research Centre, University of Turku, FI-20014 Turku, Finland
e-mail: hnurmi@utu.fi

T. Meskanen
Department of Mathematics and Statistics, University of Turku, FI-20014 Turku, Finland
e-mail: tommes@utu.fi

A. Pajala
Public Choice Research Centre, University of Turku, FI-20014 Turku, Finland
e-mail: anpaja@utu.fi

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*, DOI: 10.1007/978-3-642-35929-3_26,

Fig. 1 External and decision making cost functions



Here the horizontal axis denotes the required majority and the vertical one the costs for the individual. E denotes the external costs and D the decision making costs. The cost curves express the individual's subjective estimates of the costs as function of the decision rule. A rational individual would support that decision rule x at which the sum of these two cost types is at minimum.

In this chapter we shall utilize Buchanan and Tullock's calculus to the study of the council of ministers of European Union. As is well-known this is a weighted voting body where the weights assigned to various members states reflect—albeit in a rather crude way [1]—their population size. Over the past decades the council has been making decisions using either unanimity or qualified majority rules. Historically, unanimity has been the predominant rule, but with the Single European Act and Maastricht Treaty the qualified majority rule has been extended to cover all but the most fundamental issues (e.g. enlargement). The fact that not all decisions are made by unanimity has made the council a particularly suitable and popular case for applying *a priori* voting power measures.

The literature on the *a priori* voting power of the member states in the council of ministers is vast (see e.g. Berg et al. 1993; Brams and Affuso 1976, 1985; Herne and Nurmi 1993; Lane et al. 1996; Widgren 1994). The previous work has predominantly resorted to Shapley-Shubik and Banzhaf indices (see Banzhaf 1965; Shapley and Shubik 1954). Both indices pay explicit attention to an actor's resources (e.g. votes, seats, shares of stock) and the decision rule in the determination of his/her (hereinafter his) power measure. In addition to these two factors both indices assume that the distribution of the resources over the other actors plays a role in the computation of the power index values. In other words, the information that in a collective voting body where the total number of votes is

[1] E.g. Germany with a population of almost 80 million has 10 votes which is exactly the amount of votes that France with less than 57 million inhabitants has. Also Belgium with population size of less than 10 million has 5 votes as have the Netherlands with a 50 % larger population size.

N the decision rule is k and the number of votes of actor A is v_A is not sufficient to determine A's Shapley-Shubik or Banzhaf index value. In addition, one needs to know the distribution of votes over all the actors.

Both Shapley-Shubik and Banzhaf index measure *a priori* voting power by finding out how crucial an actor is in various coalitions of actors. Let the total number of actors be n. An actor A is crucial in coalition S if and only if his presence in S makes it winning, whereas his absence makes $S \setminus A$ non-winning. The pair $S, S \setminus A$ is then called a swing for A in S. Both indices take the swings of each player as the point of departure in determining their power index values. In fact the only difference between the indices is the weight assigned to the swings.

The standardized Banzhaf index value of A is simply the number of A's swings divided by the sum of the number of swings of all actors. A's standardized Banzhaf index value is, thus, his share of all swings in the voting body. The absolute Banzhaf index value of A is obtained by dividing the number of A's swings by 2^{n-1}, or by giving each swing an identical weight, *viz.* 2^{1-n}, and adding up the weights.

In computing the Shapley-Shubik index value of A one assigns each swing of A in S the weight $(s-1)!(n-s)!/n!$ where s is the number of actors in S. The Shapley-Shubik index value of A is obtained by adding up all weights given to A in different swings. The differences between indices thus boil down to the weights assigned to swings.

More formally, the *a priori* voting power indices are defined for simple voting games as follows. Let $v(S)$ be the characteristic function of the game, i.e. for any coalition S the function indicates the value of S. In simple voting games the value of a coalition can only be 0 or 1 with the intuitively obvious interpretation that only winning coalitions have value and each one of them has the same value, *viz.* 1. The computation formulae of the indices for actor i are then:

- the standardized Banzhaf index:

$$\bar{\beta}_i = \frac{\Sigma_{S \subseteq N}[v(S) - v(S \setminus \{i\})]}{\Sigma_{j \in N} \Sigma_{S \subseteq N}[v(S) - v(S \setminus \{j\})]}.$$

- the absolute Banzhaf index:

$$\beta_i = \frac{\Sigma_{S \subseteq N}[v(S) - v(S \setminus \{i\})]}{2^{n-1}}.$$

- the Shapley-Shubik index:

$$\phi_i = \Sigma_{S \subseteq N} \frac{(s-1)!(n-s)!}{n!}[v(S) - v(S \setminus \{i\})].$$

For easy reference the Shapley-Shubik and Banzhaf index values of various countries in the current (1996) council of ministers are presented in Table 1 assuming that the qualified majority rule is used (i.e. 62 votes out of 87).

Table 1 Shapley-Shubik and Banzhaf index values in the EU council of ministers for decision rule 62

Country	Number of votes	Shapley-Shubik index	Absolute Banzhaf index	Standardized Banzhaf index
France	10	0.1167	0.1129	0.1116
Germany	10	0.1167	0.1129	0.1116
Italy	10	0.1167	0.1129	0.1116
U.K.	10	0.1167	0.1129	0.1116
Spain	8	0.0955	0.0934	0.0924
Belgium	5	0.0552	0.0594	0.0587
Greece	5	0.0552	0.0594	0.0587
Holland	5	0.0552	0.0594	0.0587
Portugal	5	0.0552	0.0594	0.0587
Austria	4	0.0454	0.0484	0.0479
Sweden	4	0.0454	0.0484	0.0479
Denmark	3	0.0353	0.0363	0.0359
Finland	3	0.0353	0.0363	0.0359
Ireland	3	0.0353	0.0363	0.0359
Luxembourg	2	0.0207	0.0229	0.0226

2 Minimizing the External Costs

In Buchanan and Tullock's calculus the individual is assumed to minimize the costs imposed upon him by the actions of others. Viewing the EU as a collectivity of states and the council of ministers as a body controlling at least partially the actions that the states may take *vis-a-vis* each other, we may ask what decision rules would rational actors (states) endorse in an effort to curtail the external costs stemming from the actions of others. In answering this question a measure of an actor's power to prevent collective action would seem helpful. What a rational actor would, then, be envisaged to strive for is to maximize his preventive power.

2.1 The Coleman Preventive Power Index

A measure for the power to block has been defined by Coleman (1971). Actor i's power to prevent collective action is computed as follows:

$$\gamma_i = \frac{\Sigma_{S\subseteq N}[v(S) - v(S \setminus \{i\})]}{\Sigma_{S\subseteq N} v(S)}.$$

The numerator, thus, counts the swings of i, while the denominator indicates the number of winning coalitions. Accordingly, Coleman's preventive power index gives the relative number of critical presences of i in winning coalitions. Obviously, if i is a crucial member in every winning coalition, γ_i assumes its maximum

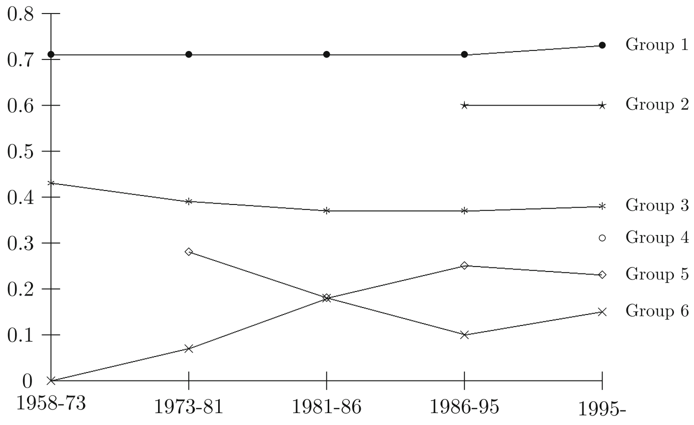


Fig. 2 Coleman's preventive power for qualified majority (71 %)

value 1. On the other hand, if i is a redundant member (dummy) in every winning coalition, $\gamma_i = 0$. It is worth noticing, however, that, in contradistinction to $\bar{\beta}_i$ and ϕ_i, it is in general not the case that $\Sigma_{i \in N} \gamma_i = 1$.

In analysing the EU council of ministers we shall adopt the following abbreviations. Group 1 includes the countries that have the largest number of votes in the council, i.e. currently France, Germany, Italy and United Kingdom. Before 1973 this group included only the first three countries. Group 2 consists of Spain, group 3 of Belgium, Greece, Holland and Portugal, group 4 of Austria and Sweden, group 5 of Denmark, Finland and Ireland, while group 6 includes only Luxembourg.

Figure 2 depicts the preventive power index values of groups 1–6 over the entire history of EU and its immediate precursors assuming that the decision rule 71 % is used in all decisions.[2] For Group 1 and 3 countries the value has remained remarkably stable. To give an idea of what the switch to simple majority rule would imply, we have computed in Fig. 3 the preventive power values for various country groups. The switch would have meant some decline in preventive power for Group 1 countries, while Group 3's power index value would have remained rather stable. A look at Fig. 2 again reveals that in "absolute" terms the simple majority rule would entail a considerable loss of preventive power for Group 1 and 3 countries. In fact, no group would benefit in terms of preventive power from the replacement of the 71 % rule with the simple majority one.

Thus, if the states are rational actors and attempt to maximize the external costs of decision making, we cannot expect great enthusiasm for moving towards the

[2] The qualified majority rule has varied between 70–71 % of all votes throughout the history of EC-EU. The exact value of the rule is always presented as the number of votes, e.g. 12 out of 17, which explains the fact that the percentage value usually has a fractional part.

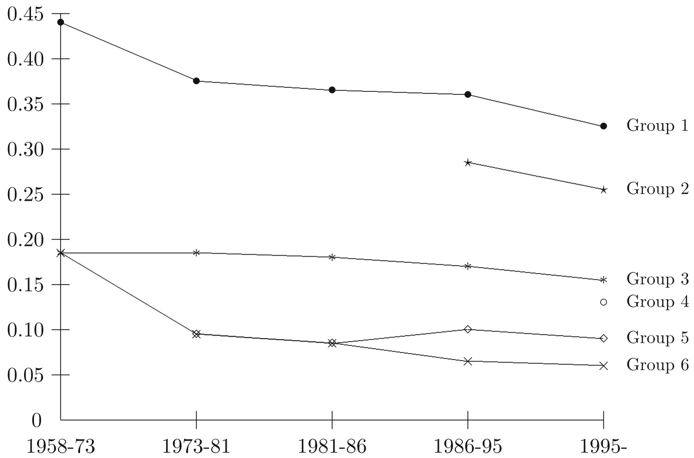


Fig. 3 Preventive power with simple majority rule

simple majority decisions in the council of ministers. In fact, even the current qualified majority rule is suboptimal in this particular sense since the preventive power would, of course, be maximized by the unanimity rule.

2.2 The Holler Index and the Johnston Index

In a recent article Johnston approaches the preventive power issue by analysing the blocking power of various countries (Johnston 1995). Instead of counting swings of various players in the voting game, Johnston takes into account only swings in a proper subset of all coalitions, *viz.* the set of minimal winning coalitions. These coalitions are defined by the property that the removal of *any* of the coalition members would make the coalition non-winning. It will be recalled that in Banzhaf indices the coalitions considered are those which are winning, but not necessarily minimal.

To illustrate the difference between Banzhaf's and Johnston's indices, consider the following voting game:

$$[4; 4, 3, 2, 1].$$

As usual the first element refers to the decision rule and the elements after semicolon denote the voting weights of the actors. Thus, 4 votes are required to pass a motion and the total number of votes is 10. Now, in this game one of the swings for the actor with 4 votes is

$$(4, 2, 1; 2, 1)$$

since the coalition of actors listed on the left side of the semicolon is winning, whereas the coalition of actors on the right side is not, and the only difference between those two coalitions is that the actor with 4 votes is present in the former but not in the latter. Thus, the above pair of coalitions is, indeed, a swing for the actor with 4 votes. Consequently, in the computation of the Banzhaf index value (absolute or standardized) of this actor the above pair is included as one swing. However, in the computation of Johnston's index value of the same actor, the above swing is not counted since the former coalition is not minimal winning: the removal of either the 2- vote or 1-vote actor or both would not make the coalition non- winning. Let us denote the set of minimal winning coalitions by $\mathcal{M}$ and the number of members in coalition S by s.

The Johnston index value of actor i, denoted $\bar{\imath}_i$, is computed as follows:

$$\bar{\imath}_i = \frac{\Sigma_{S\in\mathcal{M}} 1/s[v(S) - v(S \setminus \{i\})]}{\Sigma_{j\in N}\Sigma_{S\in\mathcal{M}} 1/s[v(S) - v(S \setminus \{j\})]}.$$

The background of Johnston's article is the debate within the EU institutions and in both houses of the British Parliament about the proper size of the blocking majority. The alternative blocking minority sizes were 23 and 27 out of the total of 90 votes. These alternatives were based on the assumption—which later on turned out to be partially incorrect—that Austria, Finland, Norway and Sweden would become new members. Johnston computes the sets of values of two indices for 16 countries assuming that the blocking minority sizes range from 23 to 27. One of the indices—which we shall call the Johnston index—is the one just defined. To the other index we shall return shortly. Johnston wants to find out whether the British strong negotiation stance in favour of the blocking minority size of 23 makes sense in terms of power indices.

It is important to notice that Johnston's index counts the swings of actors in minimal winning coalitions only, i.e. in coalitions where each member is crucial in the sense that should any member be absent from the coalition, it would no longer be winning. Johnston index, moreover, differs from Banzhaf's in weighting swings by the inverse of the number of members in the respective coalitions.

The other index that Johnston uses and somewhat unfortunately calls Banzhaf index differs from Johnston's index in assigning each swing in a minimal winning coalition the same weight. To set the record straight we shall call this index the Holler index since it was introduced by Manfred Holler in early 1980's (Holler 1982). Later it was axiomatized by Holler and Packel (1983). The Holler index value of actor i, denoted H_i, is calculated as follows:

$$H_i = \frac{\Sigma_{S\in\mathcal{M}}[v(S) - v(S \setminus \{i\})]}{\Sigma_{j\in N}\Sigma_{S\in\mathcal{M}}[v(S) - v(S \setminus \{j\})]}.$$

In Tables 2 and 3 we present the (standardized) Banzhaf, Johnston, Holler and Shapley-Shubik index values for the current 15 member council of ministers for decision rules 23 and 27. The table confirms Johnston's observation that from power maximizing point of view the stance of the UK negotiators seems to have been

Table 2 Shapley-Shubik, Banzhaf, Johnston and Holler index values in 15-member EU council of ministers for decision rule 23

Country group	Shapley-Shubik index	Standardized Banzhaf index	Johnston index	Holler index
10-vote countries	0.1206	0.1101	0.0471	0.0405
8-vote countries	0.0936	0.0927	0.0601	0.0565
5-vote countries	0.0566	0.0602	0.0716	0.0725
4-vote countries	0.0398	0.0455	0.0787	0.0820
3-vote countries	0.0332	0.0369	0.0814	0.0860
2-vote countries	0.0185	0.0248	0.0637	0.0693

Table 3 Shapley-Shubik, Banzhaf, Johnston and Holler index values in 15-member EU council of ministers for decision rule 27

Country group	Shapley-Shubik index	Standardized Banzhaf index	Johnston index	Holler index
10-vote countries	0.1191	0.1127	0.0545	0.0488
8-vote countries	0.0917	0.0911	0.0577	0.0561
5-vote countries	0.0558	0.0586	0.0728	0.0733
4-vote countries	0.0464	0.0480	0.0728	0.0751
3-vote countries	0.0313	0.0343	0.0751	0.0796
2-vote countries	0.0218	0.0246	0.0616	0.0664

irrational: to set the minority threshold at 23 instead of 27 would give UK less *a priori* voting power in terms of Banzhaf's, Holler's and Johnston's index. However, these observations do not hold if the *a priori* voting power is measured in terms of the Shapley-Shubik index. Both in the hypothetical 16-member council and in the present one 10-vote countries lose if the size of blocking minority is 27 instead of 23.

The preventive power is, however, but one of the issues that enter Buchanan and Tullock's decision rule calculus. We now turn to other considerations.

3 Minimizing the Decision Making Costs

The collective action may also have beneficial effects to the individual. Therefore, the rational individual tries to secure as much of the benefits as possible, *ceteris paribus*. In Buchanan and Tullock's book this is expressed as striving for minimizing the decision making costs. Obviously, if an individual wants the collectivity to engage in a certain type of collective action, he has to convince a sufficient number of others to vote for the action. What, in turn, is the sufficient number depends on the decision rule being applied. *A priori* it is plausible to assume with Buchanan and Tullock that the larger the majority required, the more decision making costs are involved for an individual desiring the collective action.

Table 4 Shapley-Shubik, Banzhaf, Johnston and Holler index values in the current EU council of ministers for decision rule 60

Country group	Shapley-Shubik index	Standardized Banzhaf index	Johnston index	Holler index
10-vote countries	0.1203	0.1136	0.0816	0.0800
8-vote countries	0.0939	0.0930	0.0747	0.0738
5-vote countries	0.0563	0.0584	0.0643	0.0645
4-vote countries	0.0404	0.0444	0.0606	0.0613
3-vote countries	0.0329	0.0355	0.0581	0.0593
2-vote countries	0.0202	0.0238	0.0461	0.0476

Table 5 Shapley-Shubik, Banzhaf, Johnston and Holler index values in the current EU council of ministers for decision rule 61

Country group	Shapley-Shubik index	Standardized Banzhaf index	Johnston index	Holler index
10-vote countries	0.1191	0.1127	0.0816	0.0802
8-vote countries	0.0917	0.0911	0.0741	0.0732
5-vote countries	0.0558	0.0586	0.0645	0.0647
4-vote countries	0.0464	0.0480	0.0613	0.0620
3-vote countries	0.0313	0.0343	0.0573	0.0582
2-vote countries	0.0218	0.0246	0.0474	0.0485

One way of attempting to measure the decision making costs is to employ Coleman's power to initiate index (Coleman 1971). It turns out, however, to be identical with the (standardized) Banzhaf index. For this reason the latter is sometimes called the Banzhaf-Coleman power index. Since the Banzhaf index values of the current EU member countries have already been reported for the 71 % majority rule (Table 1), we shall report in Tables 4, 5, 6, 7 and 8 the power index values of the countries for decision rules varying from 60 to 64 in the current council.

A glance at Tables 4, 5, 6, 7 and 8 suggests that perhaps the UK negotiators were not irrational, after all, in advocating the 23 minority threshold, since the local maximum power of the 10-vote countries is at 63. In other words, if those negotiators had been trying to minimize the decision making costs in the sense of Banzhaf and Coleman, the rational thing to do would have been to propose the 23 minority threshold.

The previous figures and tables reveal that in looking for the optimal decision rules the choice of the *a priori* voting power index, in terms of which one optimizes, is significant. Not only do different indices have different maxima over the domains investigated, but within the same decision making context—i.e. with a fixed allocation of votes and for a fixed decision rule—the index values exhibit interesting differences. In this regard the marked difference is between the Shapley-Shubik and Banzhaf indices, on the one hand, and the Johnston and Holler indices, on the other.

Table 6 Shapley-Shubik, Banzhaf, Johnston and Holler index values in the current EU council of ministers for decision rule 62

Country group	Shapley-Shubik index	Standardized Banzhaf index	Johnston index	Holler index
10-vote countries	0.1167	0.1116	0.0822	0.0809
8-vote countries	0.0955	0.0924	0.0751	0.0743
5-vote countries	0.0552	0.0587	0.0647	0.0650
4-vote countries	0.0454	0.0479	0.0608	0.0613
3-vote countries	0.0353	0.0359	0.0572	0.0582
2-vote countries	0.0207	0.0226	0.0440	0.0450

Table 7 Shapley-Shubik, Banzhaf, Johnston and Holler index values in the current EU council of ministers for decision rule 63

Country	Shapley-Shubik index	Standardized Banzhaf index	Johnston index	Holler index
10-vote countries	0.1197	0.1115	0.0827	0.0813
8-vote countries	0.0924	0.0926	0.0757	0.0748
5-vote countries	0.0566	0.0594	0.0645	0.0647
4-vote countries	0.0402	0.0456	0.0613	0.0620
3-vote countries	0.0331	0.0367	0.0567	0.0577
2-vote countries	0.0226	0.0228	0.0427	0.0442

Table 8 Shapley-Shubik, Banzhaf, Johnston and Holler index values in the current EU council of ministers for decision rule 64

Country	Shapley-Shubik index	Standardized Banzhaf index	Johnston index	Holler index
10-vote countries	0.1189	0.1099	0.0822	0.0808
8-vote countries	0.0884	0.0910	0.0756	0.0748
5-vote countries	0.0556	0.0591	0.0643	0.0644
4-vote countries	0.0490	0.0508	0.0611	0.0616
3-vote countries	0.0306	0.0353	0.0564	0.0576
2-vote countries	0.0237	0.0254	0.0465	0.0482

While using the former class of indices it is never the case that within the same decision making context an actor with fewer votes would have more voting power, this is possible when either Johnston or Holler indices are used.[3]

[3] While this observation cannot be made in Tables 4, 5, 6, 7 and 8, it appears if one computes the power index values for decision rules that are smaller than a simple majority. A case in point is rule 23.

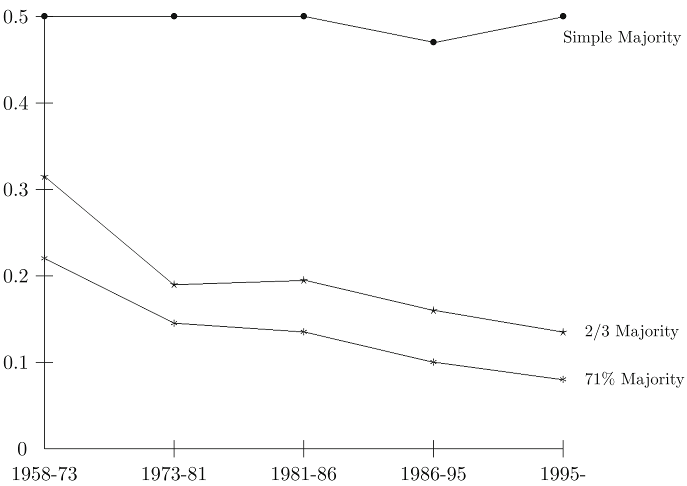


Fig. 4 The power of the Council of act for three decision rules

4 The Power of the Voting Body as a Whole

An intuitively plausible way of measuring the decision making costs in a voting body is to focus on how difficult it is to find a decisive majority on an issue. This aspect is measured by Coleman's other power index, *viz*. the power of the collectivity to act. In contrast to previous indices it is not defined for individual actors but for the voting body as a whole. Obviously, the power of the collectivity to act is a very important property since it reflects the collective capability of the body. It is, thus, related to both components of Buchanan and Tullock's calculus. If one wants to minimize the external costs of collective decisions, one should support decision rules that diminish the collectivity's power to act. On the other hand, if one is interested in getting things done in the name of collectivity, then one should propose decision rules that enhance the power to act.

Coleman's index is computed simply as the share of winning coalitions in the set of all coalitions. Formally:

$$A = \frac{\mid \mathcal{W} \mid}{2^n}.$$

Here $\mathcal{W}$ denotes the set of winning coalitions and $\mid \mathcal{W} \mid$ the cardinality of $\mathcal{W}$. Obviously, for any fixed number of actors, this index reaches its minimum value when unanimity rule is being applied.

Figure 4 shows the evolution of the power of the council of ministers to act over the history of EU for simple majority, 2/3 and 71 % majority rules.

In terms of both 71 % and 2/3 majority rules, the conclusion is clear: the power of the council of ministers to act has diminished over time. This is intuitively plausible since the addition of new members is often regarded as an obstacle to the efficient functioning of the EU machinery. It is, however, worth observing that the power of the collectivity to act would not be diminished if the simple majority rule were applied.

5 Concentration and Fragmentation

The enlargement of EU may lead to a fractionalization of the decision making system. This, at any rate, has been argued by those who support deepening rather than extending the EU cooperation. But it is well-known from the comparative studies of party systems that the sheer number of organizations called parties is not an appropriate measure of the fractionalization of the system. In search for more relevant measures, several proposals have been made. We shall discuss only two of them, one suggested by Theil and the other by Laakso (see Theil 1967; Laakso 1977; Laakso and Taagepera 1979).

Let p_i denote the share of votes of member i in a voting body of n members. In other words,

$$p_i = v_i/V,$$

where v_i is the number of votes of i and V is the total number of votes. Theil's index of fractionalization, denoted by τ, uses the auxiliary concept of entropy H which is defined as follows:

$$H = -\Sigma_i p_i \cdot \ln p_i.$$

The fractionalization index is, then, computed as:

$$\tau = e^H.$$

The minimum value of this index is 1 which results from one i having all the votes. The maximum, on the other hand, is obtained when each $p_i = 1/n$ since in that case

$$H = -(1/n)(\Sigma_i \ln 1/n) = -(1/n)(\Sigma_i - \ln n) = \ln n.$$

Thus, at the maximum:

$$\tau = e^H = e^{\ln n} = n.$$

The τ index, accordingly, equals the number of members if each member has an equal share of votes.

Laakso's index λ, in turn, is defined as:

$$\lambda = \frac{1}{\Sigma_i p_i^2}.$$

Table 9 τ-, λ- and HH-indices for the council of ministers

Index	1958–73	1973–81	1981–86	1986–95	1995–
τ	5.42	7.83	8.77	10.71	13.32
λ	5.07	7.13	7.99	9.86	12.07
HH	0.197	0.140	0.125	0.101	0.083

Again the minimum value 1 is achieved when one member has all the votes. Also the maximum value is the same for both τ and λ indices. This can be seen by assuming again that each $p_i = 1/n$ whereupon we get:

$$\lambda = \frac{1}{\Sigma_i 1/n^2} = \frac{1}{n/n^2} = n.$$

Table 9 exhibits the fragmentation history of EU council of ministers in terms of the changes in τ and λ index values.

The differences between the two indices are small and almost constant over time. It is quite obvious that the effective number of members in the EU council of ministers has always been smaller than the total number of member states. The picture we get is that of a relatively fractionalized voting body. From the view-point of EU enlargements, it is worth noticing that the addition of three small member states from the beginning of 1995 increased the effective number almost by three. In fact, the change in the effective number of countries has throughout reflected pretty closely the number of countries joining the EU being, of course, always slightly smaller than the latter. Let us now take a look at the council from the opposite angle and consider the degree of concentration exhibited by the council.

If the fragmentation is the main concern of those with a federalist turn of mind, the concentration of power in few countries is a major preoccupation of those with more critical attitude towards the integration process. Obviously, fragmentation and concentration are complementary phenomena: the more fragmentation, the less concentration and *vice versa*. Power concentration within the council can be measured by several indices. Perhaps the best-known concentration index, *viz*. the Herfindal-Hirschman index, HH, can be expressed simply as follows:

$$HH = \Sigma_i p_i^2.$$

Clearly HH is the reciprocal of λ. The HH-index values are reported in Table 9. With the increase in the number of members the concentration has been decreasing throughout the history of EU.

6 Enlargement Scenarios

In this section we shall take a look at the effects of various EU enlargement scenarios on voting power distribution among member states in the council as well as on the power of the council to act. Our first scenario is based on the assumption

Table 10 Shapley-Shubik, Banzhaf, Johnston and Coleman's preventive power index values in the 18-member EU council of ministers for decision rule 66

Country	Number of votes	Shapley-Shubik index	Standardized Banzhaf index	Johnston index	Coleman's preventive power index
France	10	0.1098	0.1031	0.0696	0.7234
Germany	10	0.1098	0.1031	0.0696	0.7234
Italy	10	0.1098	0.1031	0.0696	0.7234
U.K.	10	0.1098	0.1031	0.0696	0.7234
Spain	8	0.0845	0.0844	0.0636	0.5925
Belgium	5	0.0518	0.0543	0.0550	0.3809
Greece	5	0.0518	0.0543	0.0550	0.3809
Holland	5	0.0518	0.0543	0.0550	0.3809
Portugal	5	0.0518	0.0543	0.0550	0.3809
Austria	4	0.0435	0.0449	0.0534	0.3148
Sweden	4	0.0435	0.0449	0.0534	0.3148
Denmark	3	0.0298	0.0324	0.0490	0.2275
Finland	3	0.0298	0.0324	0.0490	0.2275
Ireland	3	0.0298	0.0324	0.0490	0.2275
Norway	3	0.0298	0.0324	0.0490	0.2275
Luxembourg	2	0.0210	0.0223	0.0451	0.1566
Cyprus	2	0.0210	0.0223	0.0451	0.1566
Malta	2	0.0210	0.0223	0.0451	0.1566

Power of the collectivity to act: 0.0742

that Norway, Malta and Cyprus will become members and that Norway is allocated 3 votes and the latter two members 2 votes each.[4] Table 10 reports Coleman's preventive power, Banzhaf, Shapley-Shubik and Johnston index value distribution for qualified majority 66/94.

We notice an instance of the paradox of new members (see Brams and Affuso 1985) here: the Shapley-Shubik index value of Luxembourg is slightly larger in the scenario of Table 10 than in Table 6. Less surprisingly this is also the case with regard to Johnston's index. Quite expectedly the power of the collectivity to act is somewhat smaller than in the present council: now it is 0.0778, while in the scenario it is 0.0742. The preventive power of the smallest members is larger under the scenario than in the present council: now 0.1471, in the scenario 0.1566. Luxembourg would, thus, benefit from this kind of enlargement.

Our second scenario has 20 members: in addition to those mentioned in the preceding scenario also Poland and Czech Republic, the former with 8 and the latter with 5 votes. The vote and voting power distribution under this scenario for decision rule 75/107 is presented in Table 11.

[4] The votes allocated to hypothetical new members are borrowed from Turnovec's article (Turnovec 1996).

Table 11 Shapley-Shubik, Banzhaf, Johnston and Coleman's preventive power index values in the 20-member EU council of ministers for decision rule 75

Country	Number of votes	Shapley-Shubik index	Standardized Banzhaf index	Johnston index	Coleman's preventive power index
France	10	0.0968	0.0904	0.0622	0.7052
Germany	10	0.0968	0.0904	0.0622	0.7052
Italy	10	0.0968	0.0904	0.0622	0.7052
U.K.	10	0.0968	0.0904	0.0622	0.7052
Spain	8	0.0762	0.0748	0.0572	0.5835
Poland	8	0.0762	0.0748	0.0572	0.5835
Belgium	5	0.0460	0.0478	0.0494	0.3729
Greece	5	0.0460	0.0478	0.0494	0.3729
Holland	5	0.0460	0.0478	0.0494	0.3729
Portugal	5	0.0460	0.0478	0.0494	0.3729
Czech R.	5	0.0460	0.0478	0.0494	0.3729
Austria	4	0.0346	0.0377	0.0457	0.2939
Sweden	4	0.0346	0.0377	0.0457	0.2939
Denmark	3	0.0274	0.0292	0.0445	0.2278
Finland	3	0.0274	0.0292	0.0445	0.2278
Ireland	3	0.0274	0.0292	0.0445	0.2278
Norway	3	0.0274	0.0292	0.0445	0.2278
Luxembourg	2	0.0173	0.0191	0.0402	0.1491
Cyprus	2	0.0173	0.0191	0.0402	0.1491
Malta	2	0.0173	0.0191	0.0402	0.1491

Power of the collectivity to act: 0.0614

The power of the collectivity to act decreases as expected. The preventive power of each member state decreases *vis-a-vis* the previous scenario as does the voting power measured by the Banzhaf and Shapley-Shubik indices.

Our final—maximal—scenario is described in Table 12. It has two additional member states: Hungary (5 votes) and Slovak Republic (3 votes) and the decision rule 81/115.

Perhaps the only slightly surprising feature of Table 12 is that the preventive power of the large member states increases *vis-a-vis* Table 11. Since the preventive power index values do not add up to 1, no other conclusion ought to be made, but the one that follows from the definition. To wit, the number of critical presences of 10-vote states in all winning coalitions is larger in 22-member council than in 20-member one. The decrease in the power of the collectivity to act is considerable when compared with the present or even with the 18-member council.

It should be emphasized that all scenarios presented in this section are based on the qualified majorities of roughly 70–71 %. Should the federalist currents prevail, there would probably be more issues to be decided in the parliament or in the council by a simple majority or both. What the employment of the simple majority rule in the 22-member council would mean in terms of power indices is indicated in Table 13.

Table 12 Shapley-Shubik, Banzhaf, Johnston and Coleman's preventive power index values in the 22-member EU council of ministers for decision rule 81

Country	Number of votes	Shapley-Shubik index	Standardized Banzhaf index	Johnston index	Coleman's preventive power index
France	10	0.0900	0.0834	0.0567	0.7184
Germany	10	0.0900	0.0834	0.0567	0.7184
Italy	10	0.0900	0.0834	0.0567	0.7184
U.K.	10	0.0900	0.0834	0.0567	0.7184
Spain	8	0.0702	0.0691	0.0521	0.5952
Poland	8	0.0702	0.0691	0.0521	0.5952
Belgium	5	0.0429	0.0446	0.0452	0.3844
Greece	5	0.0429	0.0446	0.0452	0.3844
Holland	5	0.0429	0.0446	0.0452	0.3844
Portugal	5	0.0429	0.0446	0.0452	0.3844
Czech R.	5	0.0429	0.0446	0.0452	0.3844
Hungary	5	0.0429	0.0446	0.0452	0.3844
Austria	4	0.0334	0.0355	0.0432	0.3054
Sweden	4	0.0334	0.0355	0.0432	0.3054
Denmark	3	0.0250	0.0271	0.0402	0.2333
Finland	3	0.0250	0.0271	0.0402	0.2333
Ireland	3	0.0250	0.0271	0.0402	0.2333
Norway	3	0.0250	0.0271	0.0402	0.2333
Slovak R.	3	0.0250	0.0271	0.0402	0.2333
Luxembourg	2	0.0168	0.0181	0.0368	0.1559
Cyprus	2	0.0168	0.0181	0.0368	0.1559
Malta	2	0.0168	0.0181	0.0368	0.1559

Power of the collectivity to act: 0.0485

In terms of the Banzhaf and Shapley-Shubik index the adoption of the simple majority rule would be make practically no difference. In terms of other indices, however, the decision rule is more significant. Especially noteworthy is the dramatic decrease in the preventive power of all members. That the power of the collectivity to act is considerably larger than in the qualified majority scenarios is hardly surprising.

The values of the fractionalization indices λ and τ as well as the concentration index HH are presented in Table 14 for the 18-, 20- and 22-member scenarios.

The enlargement scenarios are, thus, accompanied with a considerable increase in fractionalization. This is to be expected. The increase in fractionalization does not, however, mean that the power of the collectivity to act would necessarily decrease. In fact, as Table 13 shows, by adopting the simple majority rule the power to act can be increased from its present level. Thus, the argument that the enlargement would necessarily cripple the council decision making is valid only under the condition that the decision rules are fixed.

Table 13 Shapley-Shubik, Banzhaf, Johnston and Coleman's preventive power index values in the 22-member EU council of ministers for simple majority rule

Country	Number of votes	Shapley-Shubik index	Standardized Banzhaf index	Johnston index	Coleman's preventive power index
France	10	0.0898	0.0889	0.0482	0.2947
Germany	10	0.0898	0.0889	0.0482	0.2947
Italy	10	0.0898	0.0889	0.0482	0.2947
U.K.	10	0.0898	0.0889	0.0482	0.2947
Spain	8	0.0707	0.0700	0.0472	0.2320
Poland	8	0.0707	0.0700	0.0472	0.2320
Belgium	5	0.0429	0.0429	0.0457	0.1422
Greece	5	0.0429	0.0429	0.0457	0.1422
Holland	5	0.0429	0.0429	0.0457	0.1422
Portugal	5	0.0429	0.0429	0.0457	0.1422
Czech R.	5	0.0429	0.0429	0.0457	0.1422
Hungary	5	0.0429	0.0429	0.0457	0.1422
Austria	4	0.0334	0.0337	0.0450	0.1116
Sweden	4	0.0334	0.0337	0.0450	0.1116
Denmark	3	0.0254	0.0257	0.0448	0.0852
Finland	3	0.0254	0.0257	0.0448	0.0852
Ireland	3	0.0254	0.0257	0.0448	0.0852
Norway	3	0.0254	0.0257	0.0448	0.0852
Slovak R.	3	0.0254	0.0257	0.0448	0.0852
Luxembourg	2	0.0166	0.0169	0.0415	0.0560
Cyprus	2	0.0166	0.0169	0.0415	0.0560
Malta	2	0.0166	0.0169	0.0415	0.0560

The power of the collectivity to act: 0.500

Table 14 The Values of τ-, λ- and HH-indices for the council of ministers under three enlargement scenarios

Index	18 members	20 members	22 members
τ	15.51	17.45	19.28
λ	13.72	15.62	17.24
HH	0.073	0.064	0.058

7 Conclusions

By and large the above analysis of the changes in the *a priori* voting power distribution in the council of ministers corroborates what—on primarily intuitive grounds—has been suggested. Thus, for example, the enlargement of the EU *ceteris paribus* will lead to fractionalization and loss of the power of the collectivity to act. Also, with the advent of new members the power share of the previous ones will in general diminish. However, the paradox of new members may occur.

What is perhaps less obvious is that the enlargement of the EU may not diminish the preventive power of the existing members. Thus, it is not necessarily the case that those members that emphasize the external costs would necessarily be worse off with the entrance of new members. Similarly, it is not necessarily true that the power of the collectivity to act would diminish with new members. In all power indices utilized in this chapter, the decision rule plays an important role. Accordingly, much of the discrepancy between intuition and power index analysis stems from the fact that in the former the decision rules are not taken into account.

The fact that there are several power indices raises the question of which one—if any—is the right one in the sense of giving the most correct estimate of the real power distribution in a voting body. In our opinion none of them can do that since the real power in any voting body is based on several aspects which are not taken into account in any index. The most obvious of these aspects is the agenda-setting power both in the narrow sense of power to determine which alternatives are voted upon and in which order and in the more general sense of power to determine the policy dimensions or discourse within which the policy alternatives are to be found. What the existing power indices can be expected to do is to indicate the theoretical *a priori* influence on the outcomes that the members have by virtue of their votes and the decision rule.

It is well-known that the voting indices sometimes give widely differing values to the same members of a voting body. Thus, if one is interested in explaining or predicting which decision rules the members are supporting, the accounts based on rational behaviour are bound to differ according to which particular index value one thinks the members are trying to maximize. The fact that the Shapley-Shubik index value of an actor is at the maximum when the actor has to be present in every winning coalition has been seen as one of the advantages of the index *vis-a-vis* the standardized Banzhaf index (Laakso 1978). It is, indeed, a unique characteristic that the Shapley-Shubik index has. On the other hand, the view of coalition formation underlying this index has been often regarded as counterintuitive: while the swings are counted in each power index, the weights assigned to them in this index appear strange to many users. It is usually deemed an advantage of Banzhaf's indices that they are based on the equiprobability assumption concerning coalitions rather than permutations.

Intuitively it would seem to make sense to argue that the decision rule at which an actor has a maximum number of swings is a rough measure of his maximal influence. Hence, a good power index ought to have the maximum at this particular decision rule. Yet, this is not necessarily the case for either the Shapley-Shubik or standardized Banzhaf index. For example, in 15-member council the number of swings for a 10-vote country is about twice as large for decision rule 76 as for decision rule 78. Yet the Shapley-Shubik maximum is at 78. Laakso has several examples in which the standardized Banzhaf index differs from the Shapley-Shubik maximum. The reason is obvious: the standardized value is computed as the number of swings of an actor divided by the number of swings of all actors. The total number of swings of all actors plays a crucial role. Obviously, the absolute Banzhaf index—where the divisor is a constant—reaches its maximum exactly where the number of the actor's swings is largest.

But is the number of swings in the end the right thing to focus upon? Our discussion in Sect. 2.2 raises some doubts about it. On the other hand, Johnston's *per se* plausible idea of focusing on minimal coalitions in which each member is essential for the coalition to be winning leads to counterintuitive power indices, *viz*. those that we called the Johnston and Holler indices. Both are counterintuitive in giving in some situations member A a larger power index value in spite of the fact that member B has more votes. This possibility is not present in the other power indices.

Johnston's focus on minimal coalitions reminds us of the quite astonishing paradox of representation introduced by Schwartz (1995). According to the paradox it is possible in voting games characterized by a relatively general set of conditions that some players are better off with less representation than with more representation. This possibility is obviously present in games in which payoffs are determined by the Johnston or Holler indices. One of the conditions of Schwartz' result is that if a member is not present in two winning coalitions X and X' such that $X' \subset X$, then i prefers that X' is formed rather than X. Under Schwartz' conditions—of which we mentioned only one—it may be beneficial for a member to have less votes than more votes. Indeed, it may be better for a member to have no votes at all than to have some votes. This possibility contradicts the axiomatic properties of power indices as given by Allingham (1975), Nurmi (1997).

References

Allingham, M. G. (1975). Economic power and values of games. *Zeitschrift für Nationalökonomie Jounal of Economics*, *35*, 293–299.

Banzhaf, J. F, I. I. I. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review*, *19*, 317–343.

Berg, S., Mæland, R., Stenlund, H., & Lane, J.-E. (1993). Politics, economics and the measurement of power. *Scandinavian Political Studies*, *16*, 251–268.

Brams, S. J., & Affuso, P. J. (1976). Power and size: A new paradox. *Theory and Decision*, *7*, 29–56.

Brams, S. J., & Affuso, P. J. (1985). New paradoxes of voting power in the EC council of ministers. *Electoral Studies*, *4*, 135–139.

Buchanan, J. M., & Tullock, G. (1962). *Calculus of consent. Logical foundations of constitutional democracy*. Ann Arbor: University of Michigan Press.

Coleman, J. S. (1971). Control of collectivities and the power of a collectivity to act. In B. Lieberman (Ed.), *Social Choice*. New York: Gordon and Breach.

Herne, K., & Nurmi, H. (1993). The distribution of a priori voting power in the EC council of ministers and the European parliament. *Scandinavian Political Studies*, *16*, 269–284.

Holler, M. J. (1982). Forming coalitions and measuring voting power. *Political Studies*, *30*, 262–271.

Holler, M. J., & Packel, E. W. (1983). Power luck and the right index. *Zeitschrift für Nationalökonomie Journal of Economics*, *43*, 21–29.

Johnston, R. J. (1995). The conflict over qualified majority voting in the European Union council of ministers: An analysis of the UK negotiating stance using power indices. *British Journal of Political Science*, *25*, 245–288.

Laakso, M. (1977). Proportional methods of representation and fragmentation of the Finnish political system. *Politiikka*, *3*, 202–225 (in Finnish).

Laakso, M. (1978). The paradoxes of power indices. *Research Reports, Institute of Political Science, University of Helsinki, series A, N49* (in Finnish).

Laakso, H., & Taagepera, R. (1979). "Effective" number of parties. A measure with application to West Europe. *Comparative Political Studies*, *12*, 3–27.

Lane, J.-E., Mæland, R., & Berg, S. (1996). Voting power under the EU constitution. In S. Andersen & K. A. Eliassen (Eds.), *The European Union: How democratic is it?*. London, New Delhi: SAGE Publications.

Nurmi, H. (1997). Compound majority paradoxes and proportional representation. *European Journal of Political Economy*, *13*, 443–454.

Schwartz, Th. (1995). The paradox of representation. *The Journal of Politics*, *57*, 309–323.

Shapley, L. S., & Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review*, *48*, 787–792.

Theil, H. (1967). *Economics and information theory*. Amsterdam: North-Holland.

Turnovec, F. (1996). Weights and votes in European Union: Extension and institutional reform. *Prague Economic Papers*, *V*, 161–174.

Widgrén, M. (1994). Voting power in the EC and the consequences of two different enlargements. *European Economic Review*, *38*, 1153–1170.

The Creation of European Economic and Monetary Union

Madeleine O. Hosli

1 Introduction

European Economic and Monetary Union (EMU) was a decisive step in the European integration process. Its establishment constituted a new dimension in international monetary relations, affecting the global economic and financial environment for both public and private actors. After initial pressures on the Euro's exchange rate, EMU's new common currency appeared to be fairly stable and to have gained a rather strong position in global financial markets. In addition, in economic terms, EMU reinforced the integration process of the European Union (EU), as the euro had induced increased price transparency and decreased transaction costs across the EU. Nonetheless, skeptic voices regarding both the structure, and overall desirability, of EMU still existed. Recently, in the wake of the global financial crisis and the Euro sovereign debt crisis, the foundations of EMU have again come under scrutiny. It is conceivable that the multiple challenges stemming from the crisis will induce enhanced cooperation within EMU—also as regards much stricter fiscal and financial regulation and supervision. The 2012 Fiscal Compact testifies to this development.

Looking back into the history of EMU, negotiations on its foundations are interesting and reveal new aspects. Such aspects may, in fact, still be relevant today, as they show the limitations to the current framework of monetary union—and the ways in which EMU's foundations have come about, on the basis of intergovernmental negotiation and agreement.

An earlier version of this chapter has been published under the title "Negotiating European Economic and Monetary Union" in *Homo Oeconomicus* 25 (2008).

M. O. Hosli (✉)
Department of Political Science, Leiden University,
P.O. Box 9555 2300RB Leiden, The Netherlands
e-mail: hosli@fsw.leidenuniv.nl

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_27,

Why was EMU set up in the first place? How did ideas about monetary integration translate into concrete choices on the political level? Which member-state governments of the then European Community (EC) supported the creation of a monetary union and what priorities did they have regarding the specific institutional and substantive features to be incorporated into EMU?

To many observers, it seems especially paradoxical that Germany—with its Bundesbank widely viewed as a kind of hegemonic actor within the European Monetary System (EMS)—agreed to establish EMU. The potential costs of EMU to Germany, and its 'patience' regarding the institutionalization of monetary union, however, might have increased its bargaining leverage in the intergovernmental negotiations. In addition, as treaty revisions in the EC, and later the EU, always required unanimity on the intergovernmental level, and subsequent domestic ratification, the bargaining leverage of smaller and medium-sized states could have been enhanced in the EMU negotiation process.[1]

Compared to Germany and other larger EU states, what was the performance level of small and middle-sized EC states in the negotiation process? To what extent did they influence the intergovernmental negotiations and what exactly were their priorities regarding EMU?[2] How do their initial preferences compare to the final outcomes of the intergovernmental bargaining process?

Positions taken by the various government delegations on EMU are likely to have been representative of prevalent domestic societal interests.[3] However, this chapter will not study the process of the formation of government preferences regarding EMU, but rather, aim to assess and analyze government priorities as they were carried into the actual negotiation process. This study builds on insights provided earlier in Hosli (2000), but in addition, includes information on actors' preference intensities in the analysis. Finally, it will incorporate into the calculations a modification regarding measurement of delegations' preference intensities.

The chapter is structured as follows. The next section provides information on the distribution of government preferences with respect to the move from the EMS to EMU, derived from a data collection based on expert interviews. Section 3 of the chapter applies some negotiation analytic techniques to assess the extent to which the bargaining outcomes on EMU reflected the initial preferences of government delegations—particularly regarding issues such as the timing and structure of EMU—by taking information on preferences and preference intensities into account. The final section summarizes and evaluates the main findings of the chapter.

1 E.g., see Martin (1994, p. 88).

2 For a detailed account of the negotiations on EMU, see Dyson and Featherstone (1999).

3 Nonetheless, it seems that in Germany, for example, few domestic interest groups actually favored EMU (e.g., Verdun 2002).

2 Government Preferences Regarding EMU

Information on government preferences regarding various issues negotiated within the EC during the 1980s and 1990s, such as limits on transport fuel emissions, controls on radioactive contamination, and air transport liberalization, is available in the edited volume *European Community Decision Making* by Bueno de Mesquita and Stokman (1994). The book also provides information on government preferences regarding the establishment of EMU and the European Central Bank (ECB), in chapters by van den Bos (1994, p. 64) and notably, Kugler and Williams (1994, pp. 208–212).[4] Kugler and Williams provide thorough evaluations of the bargaining processes on EMU, aiming to predict bargaining outcomes on the basis of detailed information on preferences held by a range of influential actors in the EC. The analysis includes governments and supranational actors and utilizes models allowing for preference changes over time.

Van den Bos (1994) describes the ways data have been collected for the de Mesquita and Stokman volume: usually, data were derived by means of expert interviews, with some additional data having been collected on the basis of *Agence Europe*. In the case of expert data, interviewees—usually experts who had been involved in the negotiating process—were asked to locate member states' policy positions on given policy scales. It was understood that assessments would be on an interval scale with a maximum of 100: a distance between points of 60 and 80, for example, was to be equal to the distance between 80 and 100. Similarly, the distance between 10 and 30 was to be twice as large as the distance between, for example, 30 and 40. In the edited volume, bargaining on monetary integration is considered to constitute a special case, as EMU negotiations were conducted in an intergovernmental setting and hence, required unanimity among EC governments instead of a qualified majority.

Generally, governments' policy positions (used interchangeably with 'preferences' in the book), as well as the importance attributed to issues ('salience'), were assessed by the researchers *ex ante*, i.e. before the respective negotiations set in.[5] The subsequent analysis will use some of these data. In addition, it will adopt the authors' assumption that the experts interviewed were indeed able to locate actor ideal points on relevant scales and to provide information on government 'preferences' (instead of, for example, 'policy positions', which could also include actors' strategic considerations). In the Bueno de Mesquita and Stokman volume, different models were then applied in order to replicate the negotiation dynamics

---

4 Also see Hosli (2000).

5 Unfortunately, the exact timing of the expert interviews cannot easily be discerned from either Van den Bos (1994) or Kugler and Williams (1994). However, it appears that, regarding EMU, data on the preferred 'kind of banking arrangement' were collected just before the June–July 1989 European Community (EC) summit meeting in Madrid, and those on the remaining 'contending issues' (i.e. all other categories) just before the December 1991 meeting in Maastricht.

and, more specifically, to forecast equilibrium policy outcomes for a range of policy areas. Among the most prominent models in the book, compared to one another in terms of their relative predictive capabilities, are Bueno de Mesquita's expected utility model and an exchange model presented by Stokman and Van Oosten (1994). Both the exchange and the expected utility models are found to provide rather accurate forecasts of the actual negotiation outcomes.[6]

In a contribution regarding possible enhancement of measurement in the de Mesquita and Stokman data collection, Achen (1999) suggests an improvement of information on actors' issue 'salience'. As models need to be invariant to the scales of the measured quantities, data on preference intensities, according to Achen, could profitably be transformed from the original data set by raising them to the power of 3.1.[7] This transformation of salience scores will be used in this chapter along with original scores in order to provide alternative assessments.

Similar to an earlier analysis provided by Hosli (2000), this chapter will use some data contained in the de Mesquita and Stokman volume, but with a different goal in mind: of interest here is not the establishment—or refinement—of models aiming to make accurate predictions of bargaining outcomes. Rather, the study aims to assess the extent to which results of the intergovernmental negotiating process reflect the initial preferences held by relevant actors, notably the delegations of EC member states involved in the bargaining process. Hence, the current chapter, in correspondence to Hosli (2000), simply aims to establish how close actual negotiation outcomes were to the initial preferences of government delegations, in order to see whose preferences were most closely mirrored in the design and structure of EMU.

Based on data provided by Van den Bos (1994, p. 64), Table 1 summarizes EC government preferences regarding EMU, presenting original data on preferences as well as (transformed) data on actor salience (indicated simply by original scores being raised to the power of 3.1).

According to Kugler and Williams (1994), the first category of Table 1 ('kind of banking arrangement') constituted a major source of contention before the 1989 EC summit meeting in Madrid. In fact, this issue was only fully resolved in the 1991 summit in Maastricht. The other categories for which information is

---

[6] However, a transformation of the original 'salience' data, aiming to make them fully invariant to the respective measurement scale, appears to possibly increase the predictive accuracy of simpler models of decision-making compared to more sophisticated models presented in the edited volume (see Achen 1999).

[7] In the original analysis conducted in the 1994 volume, the scale for the intensity of preferences ranges from 0 to 100, with 50 indicating a 'neutral' position of an actor towards an issue. This method could be somewhat problematic within selected models, however, as the scores may then not fully reflect what they were designed to measure: for example, does 'of vital importance' (raw salience score 100) imply only twice as much impact in calculations as 'neither important nor unimportant' (raw salience score 50)? Accordingly, the top of the measurement scale may need to be 'stretched' relative to the lower sections. Hence, Achen suggests the relationship between measured salience $\hat{s}_{ij}$ on issue j and theoretical salience s_{ij} should have the form $s_{ij} = \hat{s}_{ij}^{\alpha}$, and, on the basis of a statistical exploration, estimates α to be 3.100 (Achen 1999, p. 11).

Table 1 Policy positions regarding an integrated European monetary and banking system (salience in brackets [a])

EU member state	Kind of banking arrangement	Time of institutionalization	Power over policies	Scope of responsibilities	Harmonization	ECU/National currencies
France	100 (75^{α})	100 (90^{α})	70 (90^{α})	90 (90^{α})	100 (70^{α})	40 (80^{α})
Germany	80 (100^{α})	30 (100^{α})	100 (100^{α})	65 (100^{α})	50 (70^{α})	90 (90^{α})
Italy	85 (50^{α})	80 (70^{α})	60 (50^{α})	75 (60^{α})	30 (30^{α})	70 (60^{α})
United Kingdom	50 (90^{α})	30 (60^{α})	60 (90^{α})	40 (90^{α})	30 (90^{α})	100 (50^{α})
Spain	100 (50^{α})	80 (60^{α})	60 (70^{α})	80 (50^{α})	30 (50^{α})	70 (60^{α})
Belgium	80 (50^{α})	50 (50^{α})	70 (90^{α})	65 (50^{α})	60 (50^{α})	70 (50^{α})
Greece	80 (60^{α})	50 (30^{α})	60 (40^{α})	100 (60^{α})	30 (50^{α})	70 (40^{α})
Netherlands	80 (90^{α})	50 (50^{α})	70 (60^{α})	65 (60^{α})	60 (50^{α})	60 (50^{α})
Portugal	80 (75^{α})	30 (30^{α})	60 (50^{α})	30 (70^{α})	30 (40^{α})	70 (30^{α})
Denmark	50 (60^{α})	30 (70^{α})	60 (75^{α})	30 (80^{α})	25 (50^{α})	40 (50^{α})
Ireland	60 (50^{α})	30 (25^{α})	60 (40^{α})	80 (50^{α})	30 (20^{α})	70 (40^{α})
Luxembourg	80 (40^{α})	50 (70^{α})	40 (40^{α})	50 (50^{α})	60 (40^{α})	60 (40^{α})
Mean position	77.1 (65.8^{α})	50.8 (58.8^{α})	64.2 (66.3^{α})	64.2 (67.5^{α})	44.6 (50.8^{α})	67.5 (53.3^{α})
Median position	80	50	60	65	30	70
Actual policy outcome	85	50	70	70	60	75

Source Adapted from van den Bos (1994, pp. 62–65) and Kugler and Williams (1994, pp. 208–212)

[a] "Salience" is the intensity of a government delegation's preference on an issue. Compared to the original coding, raw salience scores may be transformed by raising them to the power $\alpha = 3.1$, in accordance with an estimate by Achen (1999) on other policy issues included in the Bueno de Mesquita (1994) volume. Accordingly, 0^{α} indicates "of no importance", 50^{α} "neither important nor unimportant" and 100^{α} "of vital importance"

available in Table 1 were ‘remaining issues’, to be resolved in the framework of the final 1991 Maastricht meeting.

In this chapter, the figures presented in Table 1, and calculations with respect to bargaining ‘success’ in the intergovernmental negotiations, refer to the actual outcomes as incorporated into the text of the Maastricht Treaty (and not to later developments as regards the actual initiation of EMU).[8] In general terms, the timing of the EMU institutionalization reflected a battle between proponents of the ‘coronation strategy’ (‘fiscal convergence first’)—most importantly Germany, the Netherlands, Luxembourg, the United Kingdom–and the ‘locomotive strategy’ (‘institutionalization first’)—France, Italy and Belgium. The Delors Committee, which largely prepared the provisions for EMU[9], could not agree on a solution for this issue and essentially left it to be dealt with by the Intergovernmental Conference. The plain text of the Maastricht Treaty provisions seems to follow the coronation strategy: a common currency and a common central bank only after the strict fulfillment of the so called Maastricht criteria—that is only after fiscal convergence. In a historical perspective, however, it has to be acknowledged that in 1998, only Luxembourg strictly complied with all of the four fiscal requirements. So at the end, the “locomotive strategy” prevailed in practice, at least for the core countries of EMU, something the data collected by Kugler and Williams are of course unable to reflect. The data, however, do mirror government preferences on major EMU issues to be negotiated.

In essence, according to the authors, there were six relevant categories in the negotiations encompassing preferences for institutional arrangements and the ‘timing’ (or transition schedule) for EMU. Evidently, any classification into specific issue categories will simplify the analysis of the bargaining process. Nonetheless, the substantive contents of these categories will be adopted here and described in ways as closely as possible to the original data compilation.

The first category, ‘kind of banking arrangement’, encompasses data on a scale ranging from a preference to maintain the *status quo* in terms of monetary integration (scaled as 1 by the researchers) to the view that a truly ‘supranational’ bank should be established in a process of government-led convergence—scaled as 100. The preference for a common central bank to be established, only after some market-led convergence among EC economies had been achieved, is assigned a location of 20 on the scale, while the view that a central bank should be established at a later stage during a government-led convergence process is located at 80. On the basis of the expert interviews, the preference of the German delegation is indicated as 80 and that of France as 100. Belgium, Greece, the Netherlands, Portugal and Luxembourg are reported as agreeing with the German position (80).

[8] The following sentences closely follow comments made by a reviewer of this manuscript. I would like to thank this reviewer for insightful comments and suggestions made, including the observations described here.

[9] See Committee for the Study of Economic and Monetary Union (1989), Verdun (1999).

By comparison, Denmark and the UK preferred an outcome closer to the *status quo*, with preferences located at 50 on this (major) bargaining dimension.

Regarding preferences for the timing of EMU institutionalization, France favored an early start: the top of the scale (100), reflecting the ideal point of the French delegation, stands for institutionalization ‘at the earliest possible’ date (1992 in practice). 80, the option preferred by the Italian and Spanish delegations, represents a preference for institutionalization once the Single Market was in place. Delayed institutionalization—until changes accompanying monetary and fiscal convergence would have been completed—was scaled as 30 and advocated by the delegations of Germany, the UK, Denmark, Ireland and Portugal.

The next category in the data compilation is ‘power over policies’, reflecting preferences regarding who within EMU should design and execute monetary policy. The most contentious issue within this category was the allocation of powers to the ECB as a new institution compared to central bank governors. Positions taken by government representatives on this issue were between 40 and 100: the lowest score, 40, indicates a preference for power to be exercised exclusively by a council of national central bank governors. This council would hence set and execute monetary policy within EMU independently. The government of Luxembourg apparently preferred this option. A preference of 60 on the scale reflects the position that national central bank governors should direct monetary policies, while a European Bank Board would execute them. This option, according to the data collection, was supported by the governments of Denmark, Greece, Ireland, Italy, Portugal, Spain, and by the UK delegation. The governments of Belgium, France and the Netherlands, by comparison, advocated an option whereby the council of national central bank governors and the European Bank Board would largely share policy competencies (70). The German delegates seem to have favored an option where an independent European Bank Board would set and execute monetary policy (100).[10]

Preferences regarding the ‘scope of responsibilities’ to be attributed to the ECB are shown in column four of Table 1. From the data collection, it is evident that some governments would have liked to provide the new central bank with a wide range of policy competencies. For example, the Greek delegation advocated a right for the ECB to intervene in members’ domestic economies (scaled as 100). The French government representatives, with a preference located at 90 on this dimension, did not support these powers as strongly as Greece did. But they nonetheless favored establishing some ECB powers to intervene in the domestic economic sphere (i.e., beyond providing the ECB with the authority to control inflation by setting of common interest rates on the basis of the new common currency). Regarding the scope of responsibilities to be attributed to the ECB, the French delegation’s stance is said to have been mirrored by Ireland and Spain (80 each). The Italian delegation,

[10] This assessment corresponds with the observation made by some authors that the German government, to a certain extent, was not opposed to curtailing the powers of its own central bank. E.g. see Kennedy (1991), Wolf and Zangl (1996), or Cooper (1997).

with a score of 75, advocated an intermediate position between powers to intervene more drastically in domestic economies and a preference that the ECB should only hold the power to guide member states regarding inflation (65), an option favored by Belgium, Germany and The Netherlands. The allocation of purely ‘executive functions’ to the ECB on the other hand, was advocated by the government delegations of Denmark and Portugal (both located at 30 on this scale). The preferences of the delegations of the UK and Luxembourg on this scale were located at 40 and 50, respectively. The most radical view on this dimension—not advocated by any EU government in practice—would have been to limit the ECB’s responsibilities to the management of accounts.

Another dimension refers to the tools to be applied in order to harmonize national economies. Government positions on this issue were assessed on a scale—corresponding to preferences which ultimately materialized—beginning with the position that harmonization should not be directed collectively, but left to domestic governments to monitor (30). According to the data, this option was preferred by the governments of Greece, Ireland, Italy, Portugal, Spain and the UK.[11] With a score of 25, Denmark rather favored national autonomy on this issue and a purely market-led process of economic convergence. Located on the top of the scale on this dimension, and advocated by the French delegation, is a preference for collective and coordinated macroeconomic harmonization, by first linking the major economies and then managing the convergence of the smaller ones (100). An intermediate, and less directive, position on this scale was that major economies should be linked first, providing an example (and a possibility) for others to catch up in a manner suitable to them (60)—a position that seems to have been supported by Belgium, the Netherlands and Luxembourg. The German delegation is reported to have had a preference of 50 on this scale, being slightly more in favor of leaving responsibilities for harmonization in the hands of national authorities, compared to the position taken by the Benelux states.

The last category on which Kugler and Williams (1994) provide data concerns the relationship between the common currency—then still the ECU—and EC member states’ domestic currencies, and the goal to control currency fluctuations within the EC. Positions on this scale, as materialized in practice, range from a preference that all national currencies should be included into a new ECU basket (40), a position advocated by the delegations of Denmark and France, to the view that a ‘hard ECU’ should be established alongside existing currencies. The latter option was close in its orientation to the ‘parallel currencies approach’, according to which domestic currencies of EC states would become legal tender across the EC and with this, encourage market competition. As this proposal was tabled by the UK, not surprisingly, the UK delegation is located at 100 on this scale. Several EC states, according to the data collection, appear to have supported the proposal to replace existing currencies by one strong currency (60), possibly one modeled

[11] Somewhat surprisingly, according to the data collection, countries such as the UK and Greece, for example, held identical preferences regarding procedures for economic harmonization.

after the example of the German mark. Governments located closest to this position on the scale include Luxembourg and The Netherlands (60 each), and Belgium, Greece, Ireland, Italy, Portugal and Spain (70 each).

A special note is warranted on the strategies and position of the UK.[12] The UK had been strongly opposed to EMU institutionalization for a long time. At the point the Kugler and Williams data were collected, London, however, belonged to the—in fact later successful—'coronation group'. After the battle could not be won as regards avoidance of EMU institutionalization, London proposed two possible strategies for the creation of a common currency, based on liberal market principles: first, a parallel currencies approach which, by making existing national currencies legal tender in all EC member states, would allow the market to choose the most suitable currency to become the common European currency. Second, London proposed the 'hard ECU approach', which would have created a special hard currency along the existing national currencies in order again to allow for choice by market forces. In neither approach, however, did London envisage any collective institutionalization, let alone the creation of a common central bank. In a sense, London supported the coronation strategy only as the second best solution—in fact the one which seemed most likely to postpone institutionalization for quite some time. Of course, while some elements of the UK strategy are contained in the Kugler and Williams data, this two-step approach cannot fully be reflected by them.

In practice, when collecting data, it is far from an easy endeavor to represent concrete policy choices on 'ordered dimensions', as the example of the dimension 'kind of banking agreement' illustrates: this dimension appears to contain elements of both the timing and structure of the central banking system and hence, is difficult to represent in terms of choices on continuous scales. In addition, whereas the strategy of interviewing experts in order to discern government preferences has the advantage of a high level of information and extensive internal consistency of information (secondary sources, by contrast, may not be as directly comparable to each other), the actual involvement of experts in the negotiation processes, in turn, may also create biases in the data set. Clearly, there is no ideal way to go about facing such dilemmas. As Kugler and Williams, however, clearly were as careful as possible as regards data collection and measurement, the data will be used in full in the subsequent analysis. However, for data on preferences (and, in fact, issue salience), different scenarios will be worked with in an attempt, to the extent possible, to avoid any potential distortion in the analysis. Clearly, the Kugler and Williams data set on government preferences regarding different elements of EMU is unique and may be used rather effectively to shed light on the bargaining process that led to the institutionalization of EMU.

In addition to classifying member states' preferences on the given dimensions, data on the intensity of actors' preferences ('salience') were collected. These are also reported in Kugler and Williams (1994) and Van den Bos (1994),

[12] The following passages are again largely taken from helpful comments by a reviewer of this manuscript.

respectively. For example, the governments of both France and Germany, as shown by the data collection, had fairly strong preferences regarding the timing of EMU, although their delegations advocated rather different solutions for this issue. Similarly, regarding the scope of responsibilities to be attributed to the ECB, Germany, France and the UK strongly preferred specific policy options. But in comparison to the German delegation, the French representatives favored more, and the UK representatives fewer, policy competencies for the ECB. The relation between the new common currency and domestic currencies was of high importance to the German government,[13] whereas countries such as Greece, Ireland, Luxembourg and Portugal paid relatively little attention to this specific issue on the negotiation agenda.

The actual bargaining results on all categories—as contained in the provisions of the TEU with respect to EMU—have also been classified by the authors and located on the respective scales (see Kugler and Williams 1994, p. 206). Subsequently, the final bargaining outcome on issue j will be denoted R_j. According to the authors, the final banking arrangement settled upon in the negotiations corresponded to the vision that the ECB be established after a process of market-led convergence (scored as $R_{BA} = 85$). With respect to the timing of institutionalization, the final agreement was a compromise between delayed institutionalization (until the changes accompanying monetary convergence had occurred) and institutionalization once the single market had come into effect (i.e., $R_{TI} = 50$). With respect to the category 'power over policies', the negotiations essentially led to the solution that the ECB and the national central banks would share policy competencies ($R_{PP} = 70$), as indeed is reflected in the establishment of the European System of Central Banks (ESCB). The outcome regarding the range of powers to be attributed to the ECB is that the institution should not be able to directly interfere in domestic macroeconomic policy choices, but that it be given more than simply the power to control inflation ($R_{SR} = 70$). Regarding the dimension 'harmonization', the intergovernmental bargaining result corresponds to the policy choice 'link the major economies and let others catch up with them as they can' (i.e., $R_{HA} = 60$). Finally, for the last dimension, the negotiation results essentially reflect the preference that a common single currency should replace existing domestic currencies ($R_{CU} = 75$).

Given these policy preferences, preference intensities and final bargaining results, whose preferences were most strongly reflected in the actual negotiated outcomes? This question also guides the analysis in Hosli (2000). However, the current chapter extends that study, by also accounting for information on issue salience in the negotiation process. In the subsequent analysis, original actor positions will be compared with actual bargaining outcomes. Hence, more complex bargaining dynamics, such as challenges of players to other actors, exchanges of voting positions among governments, or adaptations of player preferences between

[13] Somewhat striking with respect to the coding, however, is that the UK appears to have been quite neutral on this issue (raw salience score 50), in spite of the fact that it had tabled the proposal for a 'competing currency approach' (a preference located at 100 on this scale).

summit meetings, are not accounted for here. By comparison, they are crucial to several models presented in the Stokman and de Mequita volume, as well as to the analysis regarding EMU provided by Kugler and Williams (1994). The subsequent exploration will, hence, be interested in a specific aspect of the overall dynamics: the comparison between original actor positions and ultimate negotiation outcomes.

3 'Value Scores' in Negotiations

Which government delegations were most 'successful' in the bargaining process leading to EMU? How, and to which extent, can this be assessed with the available data on EMU institutionalization? Intuitively, 'success' in negotiations implies that the final bargaining outcome is close to one's own initial policy preferences, notably regarding issues for which one's preference intensity is high. The aforementioned data compilation can be used, for example, to derive value structures (or 'value scores') for bargaining games, here for the governments involved in the negotiation process,[14] as suggested by Keeney and Raiffa (1991). In the perspective of Keeney and Raiffa, value scores are helpful as tools to be applied to achieve efficient bargaining results. Specifically, their suggestions aim to spell out and clarify the interests and preferences of negotiating parties, as well as the importance they attach to specific issues, in an attempt to find solutions allowing for efficient exchanges. Value scores, however, may also profitably be applied for other purposes: in the subsequent analysis, they will serve to assess, *ex post*, which negotiating parties in the EMU bargaining process achieved results closest to their original preferences, especially when also accounting for issue salience. In a first approach, values regarding the various issue categories ('dimensions') will be considered to be additive (i.e., values on the various dimensions will simply be summed up).

If the maximum attainable score for each relevant player is a value of 100 for each dimension given in columns one through six of Table 1, and the actual negotiation outcome on issue j is R_j, the value score v_{ij} for player i on issue j is

$$v_{ij} = 100 - |R_j - x_{ij}|, \tag{1}$$

where x_{ij} is actor i's ideal point on issue j. Applying Eq. (1) to all twelve EC governments involved in EMU negotiations, for all dimensions, generates the results provided in Table 2.

Assuming scores are simply additive for the six issue dimensions at stake, these calculations demonstrate that most of the then twelve EC member states did quite well regarding some provisions for EMU, but obtained results rather remote from their ideal points on other dimensions. For the government of France, for example, the final bargaining outcome was rather unfavorable regarding the timing of EMU

[14] This technique is also applied in Hosli (2000) based on the helpful approaches suggested by Keeney and Raiffa; for more insights into various approaches to the analysis of negotiations, see Young (1991).

Table 2 Absolute value scores (EMU negotiations)

Provisions for the EMU/ECB EU member state	Kind of banking arrangement	Time of institutionalization	Power over policies	Scope of responsibilities	Harmonization	ECU/ National currencies
France	85	50	100	80	60	65
Germany	95	80	70	95	90	85
Italy	100	70	90	95	70	95
United		Kingdom	65	80	90	70
70	75					
Spain	85	70	90	90	70	95
Belgium	95	100	100	95	100	95
Greece	95	100	90	70	70	95
Netherlands	95	100	100	95	100	85
Portugal	95	80	90	60	70	95
Denmark	65	80	90	60	65	65
Ireland	75	80	90	90	70	95
Luxembourg	95	100	70	80	100	85
Mean score	87.1	82.5	89.2	81.7	77.9	85.8

institutionalization, but corresponded closely with the French delegation's preference regarding the distribution of policy competencies between the ECB and national central bank governors. Similarly, for the German government, the negotiation outcome regarding the overall scope of activities to be conducted by the ECB was closer to its ideal point than was the result regarding the distribution of power over policies (as German negotiators favored attributing more power to the ECB than to the representatives of national central banks). A simple comparison within each dimension demonstrates that the final bargaining result regarding the 'kind of banking arrangement' coincided with the initial ideal point of Italy, whereas the result with respect to the timing of EMU institutionalization corresponded to the initial preference of Belgium, Greece, the Netherlands, and Luxembourg. Evidently, closeness of one's ideal position to the actual bargaining result does not necessarily mirror one's 'power' in the negotiation process: regarding outcomes on these dimensions, these actors may simply have been fortunate regarding the location of their preference as compared to others.

Generally, in order to increase comparability of 'success' (not to be confused with 'power') for the various government delegations in the EMU negotiation process, and assuming additivity, value scores may be normalized. Application of this procedure results in a relative value score for player i on issue j of

$$\sigma_{\mathrm{ij}} = v_{\mathrm{ij}} \Big/ \sum\nolimits_{i=1}^{n} v_{\mathrm{ij}}.$$

When the m different issue dimensions of the bargaining process are all considered to be equally important, and independent of each other, simple additive scoring leads to the summation of respective values for each player over all issues.

This leads to an overall value score for player i. In order to facilitate comparison among the n players, a normalized score σ_i for player i over all m issues, can be calculated according to Eq. 2:

$$\sigma_i = \sum_{j=1}^{m} v_{ij} \Bigg/ \sum_{i=1}^{n} \sum_{j=1}^{m} v_{ij} \tag{2}$$

Before extending these calculations, let us assume that scores were not additive—as some issue dimensions may in fact have been interrelated—and thus adapt the respective weighting. Van den Bos (1994, p. 62) mentions that the first category—the 'kind of banking arrangement' (BA)—constituted the actual centerpiece of the negotiations, whereas the other five categories reflect components of this major issue which subsequently became controversial in the negotiations. Hence, in an adapted analysis, one might either solely focus on the first dimension or, alternatively, use a different formula to derive scores on overall bargaining performance, e.g. by weighting category one stronger than the other five categories. Subsequently, it is assumed that the first category, constituting the actual centerpiece of the negotiations, should count about 0.5 of the total, whereas the rest of the overall value is composed of the remaining categories, each having equal weight. This pattern of weighting leads to the following calculation of an overall score for player i (θ_i):

$$\theta i = 0.5\,(\sigma_{iBA}) + 0.1(\sigma_{iTI}) + 0.1(\sigma_{iPP}) + 0.1(\sigma_{iSR}) + 0.1(\sigma_{iHA}) + 0.1(\sigma_{iCU}) \tag{3}$$

Hence, this approach calculates a 'multiplicative score' according to the terminology used by Keeney and Raiffa, as dimensions two through six are weighted by the factor 0.1 each, and category one by 0.5.

A more radical approach would emphasize dimension one even more heavily in the overall calculations. In order to provide a second scenario, the subsequent analysis will weight dimension one more heavily, by attributing to it a weight of 0.75 instead of 0.5 (leading to an equal weighting of the remaining categories by 0.05). Evidently, in the margin, if weights for the first dimension (the 'kind of banking arrangement'), were to be increased even further, the overall multiplicative value scores would simply converge towards the normalized scores within this dimension. Conversely, decreasing the weight of the category BA leads to convergence of the scores towards the aggregate score for categories two through six. The subsequent analysis will aim to increase the reliability of respective results by working with different scenarios regarding the weighting of categories.

In addition to the calculations presented above, data regarding preference intensity ('salience') will be included into the analysis, as derived on the basis of Van den Bos (1994, p. 64) and Kugler and Williams (1994, pp. 208–212).[15] Actors

[15] Several models in the Bueno de Mesquita 1994 volume include information on actor salience: processes of bargaining, and especially vote-trading, are assumed to be critically determined by actors' preference intensities.

strongly interested in a specific issue are generally expected to 'fight harder' for it. Intuitively, when an actor holds a strong preference on an issue, but the negotiation outcome is a large distance from his or her initial ideal point on the issue, the actor's performance is less 'successful' than if this situation applied to a topic for which the actor cared less. Conversely, a negotiation outcome close to an actor's ideal point on an issue it considers to be salient is more favorable than if an outcome were close to its ideal point on an issue for which it cared much less. Applying this intuition to respective calculations implicitly modifies the generation of 'value scores' as calculated above, as this procedure accounts for the degree to which actors are interested in specific aspects of the negotiation package. For example, Germany appears to have had a strong interest in at least the first four dimensions of the EMU negotiations as given in Table 1, attaining the maximum obtainable score ($100^{3.1}$) for the intensity of preferences on the respective (transformed) scale. Accordingly, actors' preference intensities will be accounted for adhering to this pattern, and adapted value scores are thus derived.

Including salience into the analysis, Eq. 2 may be adapted in order to calculate an (additive) normalized value score for player ϕ_{i}, accounting for preference intensities of all actors involved in the bargaining process:

$$\phi_i = \sum_{j=1}^{m} v_{ij}\, s_{ij} \Bigg/ \sum_{i=1}^{n}\sum_{j=1}^{m} v_{ij}\, s_{ij} \tag{4}$$

Transformed salience scores can be used by applying s^{α} instead of s in the calculations. Similarly, Eq. 3 may be transformed for the calculation of a multiplicative score, $\psi_{\mathrm{I},}$ including preference intensities:

$$\psi_i \ = 0.5\,(\phi_{iBA}) + 0.1(\phi_{iTI}) + 0.1(\phi_{iPP}) + 0.1(\phi_{iSR}) + 0.1(\phi_{iHA}) + 0.1(\phi_{iCU}) \tag{5}$$

Again, in order to provide an alternative estimate, the subsequent analysis will also operate with the assumption that dimension one is weighted by a factor of 0.75 and the remaining categories by 0.05 each.

Table 3 provides total normalized value scores, for the different scenarios regarding the weighting of salience (applying both 'raw' and 'transformed' values) and the weighting of the six categories respectively, on the basis of Eqs. 2 through 5.

Comparing initial actor preferences with the final bargaining outcomes, the performance of the German government, in line with findings presented earlier (e.g., Hosli 2000), indeed appears to be quite effective, notably when evaluated on the basis of formulas that also take preference intensities ('salience') into account. This favorable result for Germany can especially be seen for assessments using transformed salience scores. Hence, results generated by the use of 'value scores' according to the procedures shown above, and using the Kugler and Williams data on initial actor ideal points, support the perspective that outcomes of the negotiation process on EMU rather closely reflected the original preferences of the German government. As Table 3 illustrates, France was somewhat less successful,

Table 3 Normalized value scores (EMU Negotiations), different assessments

Scores	Value scores accounting for preferences exclusively			Value scores accounting for preferences and preference intensities ('salience')					
				Non-transformed salience scores (s)			Transformed salience scores (s^{α})		
EU member states	Additive score (Eq. 2)	Multiplicative score (Eq. 3)	Multiplicative score (Eq. 3), alternative weighing	Additive score (Eq. 4)	Multiplicative score (Eq. 5)	Multiplicative score (Eq. 5), alternative weighing	Additive score (Eq. 4)	Multiplicative score (Eq. 5)	Multiplicative score (Eq. 5), alternative weighing
France	0.073	0.076	0.079	0.150	0.145	0.135	0.207	0.188	0.158
Germany	0.085	0.088	0.089	0.142	0.137	0.134	0.262	0.254	0.253
Italy	0.086	0.090	0.093	0.082	0.077	0.074	0.054	0.047	0.039
United Kingdom	0.074	0.070	0.066	0.084	0.080	0.077	0.109	0.113	0.113
Spain	0.083	0.082	0.082	0.086	0.085	0.084	0.055	0.049	0.043
Belgium	0.097	0.095	0.093	0.082	0.076	0.071	0.065	0.048	0.039
Greece	0.086	0.088	0.089	0.069	0.072	0.076	0.036	0.041	0.047
Netherlands	0.095	0.094	0.092	0.087	0.099	0.109	0.077	0.115	0.149
Portugal	0.081	0.085	0.088	0.056	0.071	0.085	0.040	0.061	0.082
Denmark	0.070	0.067	0.065	0.056	0.053	0.051	0.050	0.040	0.036
Ireland	0.083	0.078	0.075	0.049	0.048	0.049	0.018	0.019	0.020
Luxembourg	0.088	0.089	0.090	0.057	0.058	0.055	0.027	0.025	0.020
Total	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

but follows quite closely on the scores achieved by Germany, surprisingly even surpassing Germany when non-transformed salience scores are applied (but dimensions weighted).

It is somewhat more complicated to judge the relative ‘success’ of the UK delegation in the EMU negotiations on the basis of these tools: although the UK appears to have performed well when transformed salience scores are applied, its value scores still clearly lag behind those obtained by the German government, indicating that its initial preferences corresponded rather less to the final bargaining outcomes on EMU. Partially, this may be a result of the fact that the UK was in fact opposed to EMU institutionalization, but then, as a second-best solution, joined the ‘coronation group’. The results for the Netherlands seem to be quite favorable in general. Belgium performed reasonably well, but only as assessed on the basis of calculations that do not take salience scores into account. Generally, the negotiation outcomes appear to have been rather unfavorable for Luxembourg, however, as its original preferences—also when accounting for issue salience—were quite a distance from the final bargaining results on EMU.

When salience data, and especially transformed salience scores, are taken into account, the outcomes for the Italian government are also rather unfavorable. Generally, ‘Southern’ EC states, according to this analysis, obtained results that were not very satisfactory in comparison to their initial preferences on EMU: Spain, Portugal and Greece all received scores that are rather modest, especially in assessments taking preference intensities (‘salience’) into account. This relatively weak ‘performance’ may partially be related to the fact that these countries faced domestic publics strongly in favor of EMU[16]—and with this, in the logic of two-level games, possibly had less of a ‘bargaining leverage’ in the intergovernmental negotiation. Similarly, on the basis of figures presented in Table 3, Ireland’s preferences were at a distance from the actual bargaining outcomes. Finally, according to the various assessment methods used in this chapter, the original preferences of Denmark corresponded rather badly to the final bargaining results on EMU.

4 Conclusions

Which government preferences are reflected most accurately in the final bargaining outcomes on the provisions for EMU? Which EC member states had to concede most and which states appear to be the ‘winners’ when initial preferences are compared to the final outcomes of the intergovernmental bargaining process? In an attempt to answer these questions, this chapter applies simple calculations to data on preferences and preference intensities of EC governments regarding issues such as the institutional structure and timing of EMU and the range of policy powers to

[16] I am grateful to a reviewer of this manuscript for raising this issue.

be attributed to the new ECB, in order to evaluate the intergovernmental negotiation outcomes relative to the initial preferences of the various government delegations. The chapter does not focus on whether some EC delegations held more 'bargaining power' in the negotiations than others did—since countries might simply have been fortunate that the bargaining outcomes closely reflect their initial preferences—but it is able to demonstrate the extent to which bargaining results on EMU corresponded to the original government preferences held by EC states. Accordingly, EMU preference and salience data, not only offer interesting insights into negotiation dynamics regarding the establishment of EMU, but also help to establish, *ex post*, which governments obtained relatively favorable results in these intergovernmental bargaining procedures. Hence, the current analysis provides some empirical support for more descriptive accounts on the relative performance of governments in the intergovernmental negotiation process on EMU.

Clearly, data as collected before the actual EMU negotiations set in provide a rather unique tool for investigation that can no longer be reconstructed at this point. However, in order to avoid the potential danger of bias in the calculations, the analysis conducted in this chapter resorts to different assumptions regarding the relative importance of different dimensions in the negotiation process. In addition, it uses both original salience scores and transformed ones, and employs different assumptions about the weighting of different bargaining categories.

Applying these different assumptions to the examination of the level of correspondence between initial government preferences and actual EMU provisions, the chapter finds that the German delegation avoided having to sacrifice essential interests in the intergovernmental bargaining process on EMU. Although there was some protest in German public opinion against EMU, negotiation outcomes on this issue were close to the initial policy preferences of the German government. This result is supported by each of the different methods of assessment used in the chapter, and, significantly, by evaluation models that also take preference intensities into account.

Similarly, but to a lesser extent, the French government is found to have fared relatively well in the intergovernmental negotiations. This would indeed provide some evidence for the assumption that Germany and France constituted the driving-force of the intergovernmental bargaining process on EMU. A relatively favorable outcome can also be seen for the Netherlands, a middle-sized actor. Findings for some of the remaining EC states are somewhat less conclusive. When preference intensities are accounted for, the UK appears to have obtained results fairly close to the preferences it held before the actual intergovernmental negotiations set in. By comparison, results are less favorable according to the modes of assessment that do not take preference intensities into account.

Finally, the analysis provides support for the assumption that a selection of 'southern' EC states, including Spain, Portugal and Greece, obtained bargaining outcomes that were quite distant from their original interests regarding EMU. Italy only appears to have performed well in the negotiation process when salience scores are disregarded. Finally, the outcomes of the intergovernmental bargaining process appear to be at quite a distance from the original preferences held by some smaller EU states, notably Denmark.

Generally, this chapter aims to contribute measures to assess negotiation results. Clearly, negotiation 'success' can be analyzed on the basis of different tools and approaches, and the analysis provided here—focused on the creation of EMU—shows some possibilities to do so in practice.

In the wake of the global financial crisis and the respective challenges to the Euro, the analysis of the foundations of EMU reveals interesting aspects. Clearly, if EMU was negotiated today, member state preferences on its structure and policy competences would be likely to be different from those articulated during the original bargaining process on EMU. With this, negotiation outcomes—also those related to EMU's institutional and fiscal underpinnings—might have been different, with a likely stronger focus on fiscal policy coordination and stability.

Acknowledgments Helpful comments and feedback on earlier versions of this chapter have been provided by Bryan O'Donovan, Erik Jones, Peter Lange, Anna C.H. Little, Jonathan Strand and three anonymous reviewers.

References

Achen, C. H. (1999). *Measurement puzzles in political decision making models (mimeo).*

Bueno de Mesquita, B. (1994). Political forecasting: an expected utility method. In B. Bueno de Mesquita & F. N. Stokman (Eds.), *European community decision making* (pp. 71–104). New Haven: Yale University Press.

Bueno de Mesquita, B. & Stokman, F. N. (1994). *European community decision making*. NewHaven: Yale University Press.

Committee for the Study of Economic and Monetary Union. (1989, April 12). *Report on economic and monetary union in the European community*. Committee for the Study of EMU: Brussels.

Cooper, S. (1997, August 28–31). *Governments against independent Central Banks*: *Explaining German acceptance of EMU*. Paper presented at the Annual Meeting of the American Political Science Association, Washington, DC.

Dyson, K., & Featherstone, K. (1999). *The road to Maastricht: Negotiating economic and monetary union*. Oxford: Oxford University Press.

Hosli, M. O. (2000). The creation of the European economic and monetary union (EMU): Intergovernmental negotiations and two-level games. *Journal of European Public Policy, 7*(5), 744–766.

Keeney, R. L., & Raiffa, H. (1991). Structuring and analyzing values for multiple-issue negotiations. In H. Peyton Young (Ed.), *Negotiation analysis* (pp. 131–151). Ann Arbor: The University of Michigan Press.

Kennedy, E. (1991). *The Bundesbank. Germany's Central Bank in the international monetary system*. London: The Royal Institute of International Affairs.

Kugler, J. & Williams, J. H. P. (1994). The politics surrounding the creation of the EC Bank: The last stumbling block to integration. In B. Bueno de Mesquita & F. N. Stokman (Eds.), *European community decision making* (pp. 185–228). New Haven: Yale University Press.

Martin, L. L. (1994). International and domestic institutions in the EMU process. In B. Eichengreen & J. Frieden (Eds.), *the political economy of European monetary unification* (pp. 87–106). Westview: Boulder.

Stokman, F. N. & Van Oosten, R. V. (1994). The exchange of voting positions: an object-oriented model of policy networks. In B. Bueno de Mesquita & F. N. Stokman (Eds.), *European* community decision making (pp. 105–127). New Haven: Yale University Press.

Van den Bos, J. M. M. (1994). The policy issues analyzed. In B. Bueno de Mesquita & F. N. Stokman (Eds.), *European community decision making* (pp. 33–65). New Haven: Yale University Press.

Verdun, A. (2002). *European responses to globalization and financial market integration: perceptions of economic and monetary union in Britain, France and Germany*. Houndmills, Basingstoke: Palgrave Macmillan.

Verdun, A. (1999). The role of the delors committee in the creation of EMU: An epistemic community? *Journal of European Public Policy, 6*(2), 308–328.

Wolf, D., & Zangl, B. (1996). The European economic and monetary union: 'Two-level Games' and the formation of international institutions. *European Journal of International Relations, 2*(3), 355–393.

Young, H. P. (1991). *Negotiation analysis*. Ann Arbor: The University of Michigan Press.

Apportionment Strategies for the European Parliament

Cesarino Bertini, Gianfranco Gambarelli and Izabella Stach

1 Introduction

The apportionment of seats to incoming members to the European Parliament has always been a source of discussion and research [see for instance, Laruelle (1998)]. It was discussed during the Nice Conference in 2000 and also during the accession negotiations with the applicant countries. Thanks to the Athens Accession Treaty 2003, already ten 'new' countries, joined the European Union in 2004.

The emergence of a large number of countries joining the European Union involves a considerable increase in the total number of European Members of Parliament. Generally, the 'new' countries, having weaker economies than those of existing members, could influence decisions to the disadvantage of the current members.

The trend in the past was to take into account the size of population, to attempt to guarantee representation for major political parties of each country (e.g. Luxembourg) and avoid any reduction in the number of seats held by existing members. The democratic principle 'one man, one vote'—which means that only population should be considered—has given a negative result, because many countries (with small number of citizens) rested without votes. Already Turnovec (1997) and Mercik (1999) noticed the necessity to study the problem of how to

C. Bertini (✉) · G. Gambarelli
Department of Management, Economics and Quantitative Methods, University of Bergamo, via dei Caniani 2, 24127 Bergamo, Italy
e-mail: cbertini@unibg.it

G. Gambarelli
e-mail: gambarex@unibg.it

I. Stach
Faculty of Management, AGH University of Science and Technology, Krakow, Poland
e-mail: Istach@zarz.agh.edu.pl

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*, DOI: 10.1007/978-3-642-35929-3_28, © Springer-Verlag Berlin Heidelberg 2013

divide the seats in the European Parliament, not only on the basis of the population but also taking into account other parameters.

Our proposition is to restructure the distribution of seats for all countries using a formula which takes into consideration both populations and Gross Domestic Products (GDPs). In such a way it is possible to give a simple and concrete answer to the problem under consideration. This idea will be presented in the next section. A related paradox and a number of theoretical aspects of coalition powers will be presented in Sects. 3 and 4. The optimum solutions for individual members of the new EU will be discussed in Sect. 5. Some in-depth considerations will be given in Sect. 6.

2 A Method of Seat Distribution in EU Parliament

As it has been mentioned above, we propose a new method of distributing seats by means of a formula which takes into account both populations and GDPs. The most direct method consists in adequate weighting of these data using a convex linear combination. For instance, let populations and GDP percentages of the ith country be shown by P_i and $G_{i.}$ Lets assume the weight for the population—30 % and for GDP—70 %. In this case, the seat percentages S_i of the ith country will be $S_i = 0.3 \cdot P_i + 0.7 \cdot G_i$. In general, if k is the weighting we wish to assign to the population $(0 \leq k \leq 1)$, the resulting seat percentages are $S_i = k \cdot P_i + (1 - k) \cdot G_i$. To transform the seat percentages into the number of actual seats, a suitable rounding method can be used [for instance, Hondt's proportional system, or Hamilton's Greatest Divisors, or Gambarelli's (1999) minimax apportionment, and others].

Table 1 presents the seat distribution, varying k from 0 to 1 in 10 % steps. The source of the first and last columns was IMF (2000). The GDP column was obtained by transforming the aggregate GDPs from local currencies into US dollars. For calculation of the other columns, the data used as a starting point had a greater number of decimal places and was later rounded using the Proportional System. The underlined figures show the maximum number of seats (in %) obtainable for each country.

The values in Table 1 are shown in Fig. 1a and b in a continuous way. Figure 1a indicates countries with highest GDPs and only a few countries with low GDPs. Figure 1b shows all the other countries. As can be seen, each oblique segment represents one country. If a value k is fixed on the horizontal axis and the vertical is drawn from this point, the points of intersection between the vertical and all the different segments indicate the seats (in %) to be allocated, depending on the chosen value of k. The value of the parameter k strictly characterizes the seat distribution. In fact, if $k = 0$, the seats are assigned proportionally on the basis of the countries' economic powers, without taking into account the size of population at all. And if $k = 1$, vice versa. With $k = 0$, the values of the first column of Table 1 are on the left vertical of Fig. 1, whereas for $k = 1$ the values of the last column are shown on the right vertical.

Table 1 Seats in % depending on weight k

	only ↓GDP	Values of k									only POP↓
	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Germany (GE)	**2,435**	2,361	2,288	2,214	2,141	2,068	1,994	1,920	1,847	1,773	1,700
France (FR)	**1,638**	1,598	1,557	1,516	1,475	1,434	1,393	1,352	1,312	1,271	1,230
United Kingdom (UK)	**1,589**	1,553	1,517	1,481	1,445	1,409	1,374	1,338	1,302	1,266	1,230
Italy (IT)	**1,342**	1,328	1,314	1,299	1,285	1,271	1,257	1,243	1,228	1,214	1,200
Spain (SP)	626	646	665	684	703	723	742	761	780	800	**819**
Poland (PL)	180	242	305	367	430	492	554	617	679	742	**804**
Romania (RO)	47	89	131	173	215	257	299	341	383	425	**467**
Netherlands (NE)	**428**	418	408	398	388	378	368	358	348	338	328
Greece (GR)	138	146	154	162	170	178	186	195	203	211	**219**
Czech Rep. (CR)	63	78	93	108	123	138	154	169	184	199	**214**
Belgium (BE)	**284**	276	269	262	255	248	241	233	226	219	212
Hungary (HU)	54	69	85	101	116	132	148	163	179	194	**210**
Portugal (PR)	112	122	131	141	150	160	169	179	188	197	**207**
Sweden (SW)	**269**	261	252	244	235	227	218	210	201	192	184
Bulgaria (BU)	14	30	45	61	77	92	108	124	139	155	**171**
Austria (AU)	**239**	232	225	218	211	203	196	189	182	175	168
Slovak Rep. (SR)	21	30	39	48	57	66	76	85	94	103	**112**
Denmark (DE)	**197**	188	179	171	162	153	145	136	127	119	110
Finland (FI)	**146**	142	138	134	130	126	122	118	115	111	107
Ireland (IR)	**96**	94	93	91	89	87	85	83	82	80	78
Lithuania (LI)	12	19	25	32	38	45	51	57	64	70	**77**
Latvia (LA)	7	11	16	20	25	29	33	38	42	46	**51**
Slovenia (SL)	22	24	26	28	30	32	33	35	37	39	**41**
Estonia (ES)	6	8	11	13	16	18	20	23	25	28	**30**
Cyprus (CY)	10	11	11	12	12	13	14	14	15	15	**16**
Luxembourg (LU)	**21**	20	18	17	16	15	14	12	11	10	9
Malta (MA)	4	4	5	5	6	6	6	7	7	8	**8**
Total	10,000	10,000	10,000	10,000	10,000	10000	10,000	10,000	10,000	10,000	10,000

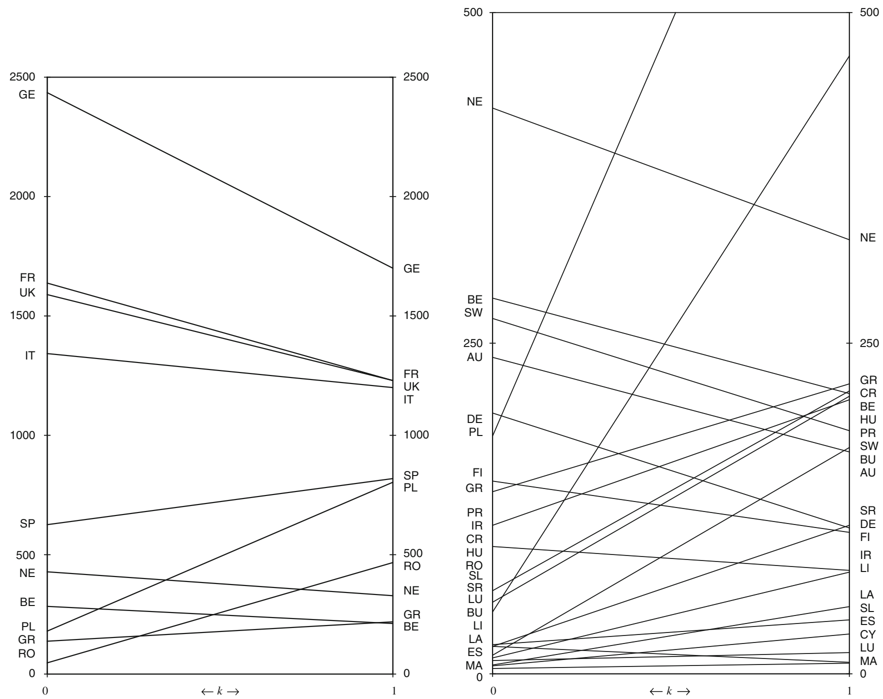


Fig. 1 **a** Seats in % of the *top* part of Table 1, depending on *k*. **b** Seats in % of *bottom* part of Table 1, depending on *k*

3 A Surprise

Once this method of apportionment is accepted, only the question of fixing the value of *k* will remain to weight the populations and GDPs, a discussion on this can be expected between countries with strong economies and those with weak economies. From an initial examination, it seems in the interests of countries with higher GDP percentages than their population percentages (Denmark, Finland, etc.) to have lower values of *k* (preferably 0), as the respective segments decrease. Conversely, for countries with lower GDP percentages (Poland, Romania, etc.) it seems advantageous to have high values of *k* (preferably 1), as the respective segments increase. However, this rule does not always apply, as will be illustrated with the following simple example.

Let us suppose that the Parliament consists of only three countries A, B and C, with relative GDP percentages of 60, 10 and 30, and population percentages of 40, 40 and 20 (see Table 2 and Fig. 2). The situation will be examined from C's point of view. It seems preferable for C, having a GDP percentage (=30) higher than its population percentage (=20), to have a very low value of *k*: preferably 0. This gives country C the maximum number of seats it could hope to attain (30 %). In this case, however, country A would obtain 60 % of the seats, i.e. the majority in

Table 2 The seat distribution in the example of three countries' parliament

	Only ↓GDP	Values of *k*					Only POP↓
	0.0	–	0.4	0.5	0.6	–	1.0
A	60	–	52	50	48	–	40
B	10	–	22	25	28	–	40
C	30	–	26	25	24	–	20
Total	100	100	100	100	100	100	100

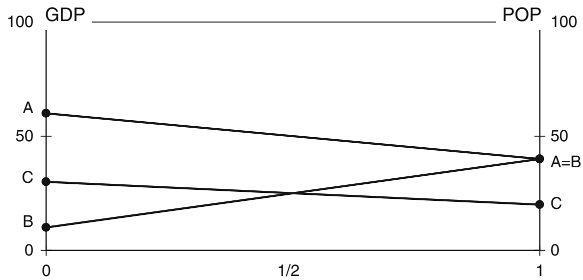


Fig. 2 The seat sharing in the example of three countries' parliament, as a function of *k*

parliament. Then, country C could not influence parliamentary decisions even if it formed a coalition with B. If, on the other hand, the maximum weight is given to the population percentage $(k = 1)$, seat apportionment is equal to that of the populations (40, 40, 20). Then the three countries have the same standing with respect to possible majority coalitions. To maintain its coalition strength, country C would be better advised to give up a certain number of seats. But how many?

4 The Power Indices

Continuing with the example above, it can be seen that C can avoid giving up a certain number of seats to maintain its coalition strength. In fact, if $k = 0.6$, seat distribution is (48, 28, 24) and C is still able to form a winning coalition with B, i.e. the sum of seats of C and B is sufficient to obtain the majority. If, on the other hand, $k = 0.5$, then the seat distribution is (50, 25, 25) and the coalition of B and C does not exceed 50 %. The latter case is of a certain interest as, with 50 % of the seats, A does not have the majority on its own and, therefore, has to form a coalition with another country. It can be seen, thus, that the 'power' of a country is not proportional to the number of its seats when we take in consideration the capacity to form a winning coalition. This power of a country is called 'coalitional power'.

It should be noticed that apportionment of seats (40, 40, 20) does not correspond with the vector of coalitional power (0.4, 0.4, 0.2) but (1/3, 1/3, 1/3) because none of the three countries has a majority on its own, and in order to obtain a majority, it has to form a coalition with another one; thus, from this point of view,

Table 3 The powers (%) corresponding to the seats from Table 2

	Only ↓GDP	Values of k					Only POP↓
	0.0	–	0.4	0.5	0.6	–	1.0
A	100	–	100	60	33.3	–	33.3
B	0	–	0	20	33.3	–	33.3
C	0	–	0	20	33.3	–	33.3
Total	100	100	100	100	100	100	100

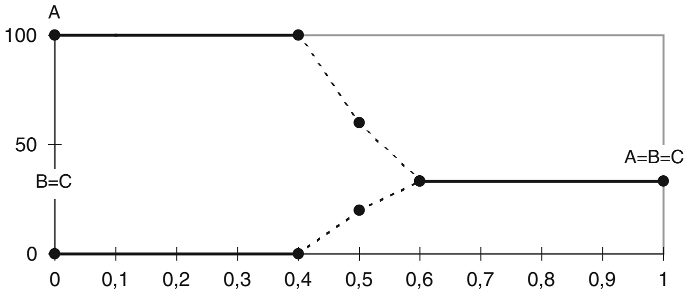


Fig. 3 The power (%) sharing in our example as a function of k

they are of equal power. We have seen that the seat apportionment (40, 40, 20) results in the division of coalitional power (1/3, 1/3, 1/3) and the apportionment (60, 10, 30) results the division of coalitional power (100 %, 0, 0).

What power can we assign to the apportionment of seats (50, 25, 25)? One method of evaluation is based on the concept of 'cruciality'. A member is said to be crucial for a coalition if this coalition is a majority with him and a minority without. In the case of (50, 25, 25) country B is crucial for just one coalition (A, B), country C for only one coalition (A, C), whereas country A is crucial for three coalitions: (A, B), (A, C) and (A, B, C). The coalitional power of each member can be expressed in proportion to the number of coalitions to which it is crucial. In our case, this power is 3/5ths (=60 %) for A and 1/5th (=20 %) for each of the other two. This power index is known as the 'Banzhaf-Coleman index' and it is one of the most applied power indices; for further explanations, please see Sect. 6.

Tables 2 and 3 and Figs. 2 and 3 present a summary of the information given so far. It can be seen that the optimum value of k for country A is 0, but the increase of k to 0.4 is not detrimental in terms of the absolute majority. With regard to B, the optimum value of k is 0.6 (which would give 28 seats) and its power would remain unchanged with a higher value of k. The optimal value of k for C is 0.6. Such a value guarantees C the maximum obtainable coalitional power (33.3 %) and the maximum number of seats (24).

An important addition must be made to complete this brief discussion of power indices. The observations assumed the formation of simple majorities (i.e. >50 %). If, on the one hand, we consider decisions with qualified majorities, the winning coalitions and, consequently, the power indices will obviously change. In decisions where there is a requirement for a majority of more than 3/4 (i.e. >75 %), the only winning coalition is (A, B, C) for the case of (50, 25, 25). Then each party is crucial only once and the index is (1/3, 1/3, 1/3). On the other hand, for the decisions requiring a majority of more than 2/3rds, the winning coalitions remain (A, B), (A, C) and (A, B, C) and therefore the index remains (3/5, 1/5, 1/5).

5 Optimum Solutions for EU Members

Table 4 gives the results of the calculations done on the apportionment of seats in Table 1. The calculations consider the threshold of a simple majority that is the prevalent rule in the European Parliament. For example, if we fix $k = 0.3$, the seats

Table 4 Powers % corresponding to the seats from Table 1

	Only ↓GDP	Values of k									Only POP↓
	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
GE	**2,856**	2,724	2,600	2,483	2,378	2,284	2,197	2,115	2,034	1,946	1,853
FR	**1,540**	1,504	1,469	1,438	1,409	1,383	1,357	1,332	1,304	1,271	1,233
UK	**1,486**	1,453	1,423	1,397	1,374	1,354	1,336	1,316	1,293	1,265	1,233
IT	1,142	1,144	1,150	1,159	1,172	1,185	1,198	1,207	**1,209**	1,207	1,199
SP	606	639	671	701	728	751	767	776	782	787	**793**
PL	179	243	308	372	435	494	552	608	662	718	**776**
RO	47	89	132	173	214	252	290	327	365	406	**452**
NE	**437**	432	421	406	389	371	356	343	332	324	317
GR	137	146	155	162	169	174	180	186	193	202	**211**
CR	63	78	93	108	122	135	149	161	175	190	**206**
BE	**283**	278	271	264	254	243	233	223	215	209	205
HU	54	69	85	101	115	129	143	156	170	186	**203**
PR	111	122	132	141	149	156	163	171	179	188	**200**
SW	**268**	262	254	245	234	223	211	201	191	184	177
BU	14	30	45	61	76	90	104	118	132	148	**165**
AU	**238**	233	227	219	210	199	189	181	173	167	162
SR	21	30	39	48	56	64	73	81	89	98	**108**
DE	**196**	189	180	171	161	150	140	130	121	114	106
FI	**145**	142	139	134	129	123	118	113	109	106	103
IR	**96**	94	93	91	88	85	82	79	78	76	75
LI	12	19	25	32	38	44	49	54	61	67	**74**
LA	7	11	16	20	25	28	32	36	40	44	**49**
SL	22	24	26	28	30	31	32	33	35	37	**39**
ES	6	8	11	13	16	18	19	22	24	27	**29**
CY	10	11	11	12	12	13	13	13	14	14	**15**
LU	**21**	20	18	17	16	15	13	11	10	10	9
MA	4	4	5	5	6	6	6	7	7	8	**8**
Total	10,000	10,000	10,000	10,000	10,000	10,000	10,000	10,000	10,000	10,000	1,0000

will be distributed in accordance with the respective column of the Table 1 (Germany 22.14 %, France 15.16 % and so on). For such division of seats, every country becomes crucial for a certain number of coalitions. The percentage of coalitions for which a given country is crucial, i.e., its power index, is presented for each country in the column '0.3' of Table 4 (Germany 24.83 %, France 14.69 %, etc.).

In Table 4, the maximum values for each country in terms of power indices are underlined. Note that Italy, although it has a GDP percentage higher than its population percentage, does not have the maximum advantage when only the GDP is taken into account (i.e., if $k = 0$). Its maximum advantage is achieved with a balanced division ($k = 0.8$). It could be seen as a paradox; for a more detailed explanation of this result, please see Sect. 4.

The figures in Table 4 provide information that may be helpful during the discussion concerning seat apportionment. The information tells each country how much the renouncement of every 1/10ths in the fixation of k costs in terms of the coalitional power. For example, the change of k from 0.7 to 0.8 costs Germany 0.81 percentage points (from 21.15 to 20.34 %), while the change of k from 0.8 to 0.9 costs the same country 0.88 % points (from 20.34 to 19.46 %). In the case of Table 1, the decrements of the percentages seats are always constant.

6 Some In-Depth Considerations

Let us follow some considerations dedicated to those readers who will want to explore in depth some of the arguments dealt with above.

6.1 On Power Indices

The power index used here was in the past known as the 'Normalized Banzhaf index' [Banzhaf (1965)] and later as 'Banzhaf-Coleman Index', thanks to the contribution of Coleman (1971). There are other indices that use crucialities, but they take it into account differently and they are based on particular bargaining models. For instance the index of Shapley and Shubik (1954) [based on the Shapley (1953) value of a game], the Holler (1978) index, and the Nucleolus [Schmeidler (1969)]. Further information on the matter can be found in Owen (1995), Gambarelli and Owen (2002, 2004), Gambarelli (1994), (1999). Some algorithms for the computing of these indices are known [see for instance, Gambarelli (1990, 1996)]. A modification of the program quoted in Bilbao (2000) and Bilbao et al. (2000) has been used for the computations of Table 4.

Various studies on the applications of power indices to the European Parliament and the Council of Ministers structure were developed in the past century [see for instance the simulations of Gambarelli and Holubiec (1990)]. More recently, the

Nice conference stimulated the emergence of very important contributions; we quote in particular Bilbao (2000, 2001), Felsenthal and Machover (1998, 2001 and 2003). Further applications based on the above quoted papers as well as Affuso and Brams (1985), Berg et al. (1996), Brams (1976), Gambarelli and Owen (1994), Garrett and Tsebelis (1999), Herne and Nurmi (1993), Holler and Widgrén (1999), Laruelle et al. (2003), Laruelle and Widgrén (1998), Owen and Shapley (1989), Turnovec (1996), Widgrén (1996) are under consideration.

6.2 On Overtaking

In the cases examined here we have varied k in steps of 1/10th, but in general this parameter can vary in the continuous interval [0, 1]. The values of k corresponding to the intersections of segments are of particular interest. In fact, each of these intersections represents a situation of equality of voting power of two countries. In formal terms, let G_i and P_i be the GDP and the populations of the ith country, respectively. An intersection between the segments representing countries i and j falls within the interval (0, 1) whenever

$$(G_i < G_j \text{ and } P_i > P_j) \quad \text{or} \quad (G_i > G_j \text{ and } P_i < P_j).$$

The corresponding value of k_{ij} is:

$$k_{ij} = (G_j - G_i)/[(P_i - P_j) + (G_j - G_i)].$$

6.3 On Optimum Weight Intervals

In the evaluations made in Sect. 4 with regard to the optimum values of k for each country, we took only variations of k with 0.1 steps into account. This led to optimum intervals [0, 0.4] for A and [0.6, 1] for B and C. If we now vary k in a continuum, we discover an interval of values even more advantageous for C. If there are 100 seats to be assigned and the Hondt Proportional System is used for rounding, then each weight k satisfying the inequalities $0.52 < k < 0.55$ leads to (49, 26, 25), which is the optimum distribution of seats for C. The optimum values of k, however, remain unchanged for the other two countries. In more theoretical terms, we can define 'optimum weight interval of the ith Country' as the variability interval k_i which guarantees this country the maximum power index, and with equal power index, the maximum number of seats. In our example it is possible to check that the optimum intervals for the three countries are as described in Table 5. Further studies on the behavior of optimum weight intervals could be carried out, based on Gambarelli (1983), Freixas and Gambarelli (1997) and Saari and Sieberg (2000).

Table 5 Optimum weight intervals for the example in Table 2

	Optimal interval	Seats	Power indices
A	(0, 0.025)	60, 10, 30	1, 0, 0
B	(0.98, 1)	40, 40, 20	1/3, 1/3, 1/3
C	(0.52, 0.55)	49, 26, 25	1/3, 1/3, 1/3

7 Conclusions

This chapter presents a proposal on how to divide seats when countries with weaker economies enter the European Union. Our proposal, considering the threshold of a simple majority, is simpler than other proposals being studied currently. This method takes into account both populations and GDP. As regards the weights to be assigned to these two components, a paradox relating to Italy was discussed. Such paradox is due to the fact that although Italy had the maximal seat percentage taking into account only GDP this country did not receive the maximal power in this situation (for more particulars see Sect. 5). Moreover, reference information for the discussion on the determination of such weights was provided for each country. Similar paradoxes, which, however, are not presented in this chapter, were also carried out with others percentages. It is a remarkable paradox relating to the Netherlands and to Ireland obtained with the percentage of 62 %.

The majority system of the European Parliament foresees majorities even distinct from the simple majority, but the authors take 50 % threshold into consideration because it seems to be the most intuitive and the choice does not have any influence on the argument presented.

A practical application of the method described above must include a proper rule for the calculation of GDP, both in the present situation and in the future updates which might be necessary, such as for example new elections, new entries, or important changes in the percentages.

The method proposed in this chapter can be a starting point for some in-depth considerations and other studies. The first necessary consideration is for the population and GDP, here considered as national items. From the high correlation existing between these two quantities, some of the arguments which hold for national population hold also for national GDP weights. It is interesting to extend the above results considering the couple—population and GDP—also in a different way which can best reflect international nature of the European Parliament where many factions of the voting body are international. In fact, in the Chamber, the members sit in political groups (PPE-DE, PSE, ELDR and so on), not in national delegations. Other consideration could be done through the analysis of the decision power of particular parliaments in the light of the studies of Owen on the a priori coalitions [Owen (1977, 1982), Carreras and Owen (1988)] and on the optimal location of candidates in ideological space [Owen and Shapley (1989)]. Last but not least consideration is the comparison of above results with the results concerning to the others power indices like those of Shapley-Shubik and Holler.

Acknowledgments This work is sponsored by MIUR. The authors would like to thank Jesus Mario Bilbao, Annalisa Cristini and Moshé Machover for their useful suggestions and Mara Grosini for help in computations.

References

Affuso, P., & Brams, S. J. (1985). New paradoxes of voting power on the EC council of ministers. *Electoral Studies, 4*(2), 135–139.

Banzhaf, J. F. (1965). Weighted voting doesn't work: a mathematical analysis. *Rutgers Law Review, 19*, 317–343.

Berg, S., Lane, J. E., & Maeland, R. (1996). Voting power under the EU constitution. In S. Andersen & K. A. Elliassen (Eds.), *The European Union: How Democratic Is It?*. London: SAGE Publications.

Bilbao, J.M. (2000), Nice European council. Retrieved December 29, 2000, from http://www.esi2.us.es/~mbilbao/niza.htm.

Bilbao, J. M. (2001). European voting games. Retrieved August 20, 2001, from http://www.esi2.us.es/~mbilbao/eugames.htm.

Bilbao, J. M., Fernández, J. R., Jiménez Losada, A., & López, J. J. (2000). Generating functions for computing power indices efficiently, *Top* 8, 2: 191–213. Retrieved August 20, 2001 from http://www.esi2.us.es/~mbilbao/pdffiles/generat.pdf.

Brams, S. J. (1976). *Paradoxes in politics: An introduction to the nonobvious in political science*. New York: Free Press.

Carreras, F., & Owen, G. (1988). Evaluation in Catalonian Parliament 1980–1984. *Mathematical Social Science, 15*, 87–92.

Coleman, J. S. (1971). Control of collectivities and the power of collectivity to act. In B. Liberman (Ed.), *Social Choice* (pp. 269–300). London: Gordon and Breach.

Felsenthal, D. S., & Machover, M. (1998). *Measurement of voting power: Theory and practice, problems and paradoxes*. Cheltenham: Edward Elgar.

Felsenthal, D.S. and Machover, M. (2001). *The Treaty of Nice and qualified majority voting, Voting Power and Procedures Report*, VPP 02/01. Retrieved August 20, 2001 from Centre for Philosophy of Natural and Social Science. London School of Economics and Political Science web site: http://www.lse.ac.uk/Depts/cpnss/projects/vp.html.

Felsenthal, D.S., & Machover, M. (2003). *Analysis of QM rules in the draft constitution for Europe proposed by the European convention, 2003*. Retrieved November 8, 2003 from Centre for Philosophy of Natural and Social Science, London School of Economics and Political Science web site: http://www.lse.ac.uk/Depts/cpnss/projects/EUcon.pdf.

Freixas, J., & Gambarelli, G. (1997). Common properties among power indices. *Control and Cybernetics, 4*, 591–603.

Gambarelli, G. (1983). Common behaviour of power indices. *International Journal of Game Theory, 12*(4), 237–244.

Gambarelli, G. (1990). A new approach for evaluating the shapley value. *Optimization, 21*(3), 445–452.

Gambarelli, G. (1994). Power indices for political and financial decision making: A review. *Annals of Operations Research, 51*, 165–173.

Gambarelli, G. (1996). Takeover algorithms. In M. Bertocchi, E. Cavalli, & S. Komlosi (Eds.), *Modelling Techniques for Financial Markets and Bank Management: the 16-th and 17-th Euro Working Group of Financial Modelling Meetings* (pp. 212–222). Heidelberg: Physica Verlag.

Gambarelli, G. (1999). Minimax apportionments. *Group Decision and Negotiation, 8*(6), 441–461.

Gambarelli, G., & Holubiec, J. (1990). Power indices and democratic apportionments. In M. Fedrizzi., & J. Kacprzyk (Eds.), *8-th Italian-Polish Symposium on Systems Analysis and Decision Support in Economics and Technology* (pp. 240–255), Warsaw: Omnitech Press.

Gambarelli, G., & Owen, G. (1994). Indirect control of corporations. *International Journal of Game Theory, 23*(4), 287–302.

Gambarelli, G., & Owen, G. (2002). Power in political and business structures. In M. J. Holler, H. Kliemt, D. Schnidtchen, & M. E. Streit (Eds.), *Power and Fairness (Jahrbuch für Neue Politische Ökonomie 20* (pp. 57–68). Tübingen: Mohr Siebeck.

Gambarelli, G., & Owen, G. (2004). The coming of game theory. In G. Gambarelli (Ed.), *Essays on Cooperative Games—in Honor of Guillermo Owen (Special Issue of Theory and Decisions)* (Vol. 36, pp. 1–18). Dordrecht: Kluwer Academic Publishers.

Garrett, G., & Tsebelis, G. (1999). Why resist the temptation to apply power indices to the EU. *Journal of Theoretical Politics, 11*, 321–331.

Herne, K., & Nurmi, H. (1993). The distribution of a priori voting power in the EC council of ministers and the European parliament. *Scandinavian Political Studies, 16*, 269–284.

Holler, M. J. (1978). A priori party power and government formation, *Munich Social Science Review* 1: 25–41. (Reprinted In M. J. Holler (Ed.), *Power*, *Voting*, *and Voting Power* (1982), Würzburg and Vienna: Physica-Verlag).

Holler, M. J., & Widgrén, M. (1999). Why power indices for assessing EU decision making? *Journal of Theoretical Politics, 11*, 291–308.

IMF (2000). International Financial Statistics Yearbook—International Monetary Fund Vol. LIII, IMF Statistics Department.

Laruelle, A. (1998). *Game Theoretical Analysis of Decision-Making Processes with Applications to the European Union*. Louvain-la-Neuve: CIACO.

Laruelle, A., Martinez, R., & Valenciano, F. (2003). On the difficulty of making decisions within the EU-25. Retrieved October 1, 2003 from http://merlin.fae.ua.es/laruelle/Download.html.

Laruelle, A., & Widgrén, M. (1998). Is the allocation of voting power among EU states fair? *Public Choice, 94*, 317–339.

Mercik, J.W. (1999), *Siła i Oczekiwania. Decyzje Grupowe*, Warszawa—Wrocław: Wydawnictwo Naukowe PWN.

Owen, G. (1977). Values of games with a priori unions. In R. Henn & O. Moeschlin (Eds.), *Essays in Mathematical Economics and Game Theory* (pp. 76–88). Berlin: Springer.

Owen, G. (1982). Modification of the Banzhaf-Coleman index for games with a priori unions. In M. J. Holler (Ed.), *Power, Voting, and Voting Power* (pp. 232–238). Würzburg: Physica.

Owen, G. (1995), *Game Theory* (3rd ed.), San Diego: Academic Press.

Owen, G., & Shapley, L. S. (1989). Optimal location of candidates in ideological space. *International Journal of Game Theory, 18*(3), 339–356.

Saari, D. G., & Sieberg, K. K. (2000). Some surprising properties of power indices. *Games and Economic Behavior, 36*, 241–263.

Schmeidler, D. (1969). The nucleolus of a characteristic function game. *SIAM Journal of Applied Mathematics, 17*, 1163–1170.

Shapley, L. S. (1953). A value for *n*-person games. In H. W. Kuhn & W. Tucker (Eds.), *Contributions to the Theory of Games (2, 39)* (pp. 307–317). Princeton: Princeton University Press.

Shapley, L. S., & Shubik, M. (1954). A method for evaluating the distributions of power in a committee system. *American Political Science Review, 48*, 787–792.

Turnovec, F. (1996). Weights and votes in European Union: Extension and institutional reform. *Prague Economic Papers* V: 161–174.

Turnovec, F. (1997, December 12–13). *Decision making games in the European Union*. Paper presented at the international workshop EU decision making reform and European integration, Prague, Czech Republic.

Widgrén, M. (1996*). A voting power analysis of supranational and national influence in the EU*. Discussion Paper 1288, Centre for Economic Policy Research, London.

Strategic A Priori Power in the European Union's Codecision Procedure Recalculated for EU28

Stefan Napel, Mika Widgrén and Alexander Mayer

1 Introduction

The question of national influence on legislation adopted by the European Union (EU) is of interest and importance to politicians, the general public, and academics alike. It has inspired a great number of applied studies and vigorous methodological debate. The applications have highly concentrated on the intra-institutional distribution of power in the EU's Council of Ministers, which is the EU's main decision-making body, using mathematical measures of voting power that have roots in cooperative game theory.[1] These studies started to mushroom in the early 1990s, and have often been inspired by EU enlargements and institutional reforms where, indeed, the Council was the key institution. A quite separate line of research has focused on inter-institutional power analysis, assuming spatial preferences and

An earlier version of this article by Stefan Napel and Mika Widgrén was published in *Homo Oeconomicus* 26(3–4), 297–316, 2009. Alexander Mayer updated the data and references, and recalculated all tables for EU28 assuming that Croatia's accession to the EU takes place on 1 July 2013 as scheduled. We thank Manfred J. Holler and Hannu Nurmi for helpful suggestions. The usual caveat applies. Mika Widgrén unexpectedly passed away on 16.8.2009 at the age of 44.

[1] For examples see, e.g., Widgrén (1994), Laruelle and Widgrén (1998), Felsenthal and Machover (2001, 2004), Leech (2002), and Baldwin and Widgrén (2004).

S. Napel (✉)
Department of Economics, University of Bayreuth, Bayreuth, Germany

S. Napel · A. Mayer
Public Choice Research Centre, University of Turku, Turku, Finland

M. Widgrén
Turku School of Economics and Public Choice Research Centre, University of Turku, 20014 Turku, Finland

A. Mayer
Department of Economics, University of Bayreuth, Bayreuth, Germany

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_29,

investigating non-cooperative voting games (see, e.g., Steunenberg 1994; Tsebelis 1994 or, for an extensive survey, Steunenberg and Selck 2006).

The cooperative index-based approach has been heavily criticized by political scientists who analyze EU decision making via spatial voting games because it does not take procedures and strategic aspects into account (see, e.g., Garrett and Tsebelis 1999). However, in the spatial voting games literature, the analysis of intra-institutional power relations, and especially the distribution of power in the Council, is still in its infancy. There, it has mostly been ignored that the Council applies a (rather complex) weighted voting rule.

In this chapter of the relative influence of the Council, we apply the unified framework for power analysis introduced in Napel and Widgrén (2004). The framework generalizes the measurement ideas underlying, e.g., the Penrose-Banzhaf or Shapley-Shubik indices to non-cooperative voting models and procedure-based strategic interaction. Thus the major limitations of traditional indices that have been pointed out by Garrett and Tsebelis (amongst many others) are overcome. The framework allows to evaluate the distribution of power at the inter-institutional and intra-institutional levels simultaneously.

We compute the distribution of power inside the Council of Ministers for a priori unknown, one-dimensional spatial preferences. A key feature, which distinguishes our intra-Council power analysis from the existing studies that we are aware of, is that we consider an actual decision procedure, namely the so-called *codecision procedure*.[2] The procedure implies that pivotality of an individual member inside the Council, which is picked up by conventional power indices, does *not* automatically translate into power to affect the collective decision. The reason is that the codecision procedure also involves the European Commission and the European Parliament; and the latter may be the truly critical player on a given issue. The individual chances of being pivotal or critical for a *decision* (rather than only for the Council's opinion on some matter), which voting power analysis ultimately is about, are affected differently for different Council members by the Parliament's presence. This means that the standard indices considered in previous investigations can give a distorted view of the actual distribution of a priori power amongst the members of the Council of Ministers.

The codecision procedure has already been investigated by Napel and Widgrén (2006) in some detail. The focus there, however, was put on the *inter*-institutional balance of power. The critical determinant of the relative influence of the Council versus the Parliament on codecision outcomes turned out to be the respective decision quotas. So, in order to simplify the analysis, we concentrated on the supermajority aspect of Council voting rules, and ignored the asymmetries in

[2] Napel and Widgrén (2011) investigated the EU's *consultation procedure*. It was introduced already in the Treaty of Rome in 1957 and for a long time remained the only way to take decisions in what is now the European Union. After the Lisbon Treaty has come into force, this procedure plays a much smaller role—mainly for competition law—than it used to. However, its relative simplicity provides an ideal framework for investigating the effects of inter-institutional interaction on voting power.

voting weight. The present paper obviously has to give up this simplification because it deals with the distribution of power inside the Council, and the weight distribution is essential for that. As a welcome side-benefit of studying weighted voting in the Council, we obtain a better assessment of the inter-institutional distribution of power than in Napel and Widgrén (2006).

The numerical differences to our earlier assessment of the power relation between Council and Parliament, and similarly to assessments of the intra-Council power distribution by standard indices, actually turn out to be relatively small. This is good news, but it should not be mistaken as an excuse for continuing with the past disregard of procedures and strategies in voting power analysis. In particular, we identify two biases of standard power measures: they count intra-institutional pivot positions for which the considered institution is outcome-irrelevant because, first, the outcome is determined by other institution(s) (here: EP) or, second, the status quo prevails because the involved institutions block each other. It may be just a coincidence that the opposite biases induced by these two types of miscountings happen to approximately cancel for most EU member states in the case of codecision under the Nice or Lisbon qualified majority rules. A change of the procedure or of these voting rules (e.g., after an enlargement beyond EU28) could easily render one of the biases dominant, and then result in much bigger deviations between non-strategic and strategic a priori power.

The remainder of the paper is organized as follows: we introduce intra-Council decision making after some preliminaries in Sect. 2. As they are our main target of assessment, some basic facts about it and the codecision procedure are needed. We then construct a simple game-theoretic model of the codecision procedure, which captures strategic inter-institutional interaction. We discuss the equilibrium outcomes predicted by this model, which are then used for the power analysis that is presented in Sect. 3. Its results are reported in Sect. 4 and, finally, we conclude in Sect. 5.

2 The Codecision Procedure

2.1 Preliminaries

In the following, we will consider the one-dimensional convex Euclidean policy space $X = [0, 1] \subset \mathbb{R}$. The legislative *status quo* regarding the (a priori random) issue which is up for a decision is denoted by $q \in X$. The considered political actors are all assumed to have *single-peaked preferences* regarding this issue. They are characterized by an individual *bliss point* or *ideal point* $\lambda \in X$: the smaller the distance $d(\lambda, x)$, the higher the agent values a policy $x \in X$. We suppose that not only do the 766 members of EP[3] and the 28 national government representatives in CM have such preferences, but that there are representatives of EP and CM who possess aggregated spatial preferences of the same kind (possibly with $\lambda = q$ if the institution cannot reach an internal consensus). It is then possible to predict the codecision outcome regarding the considered issue by specifying, first, how EP's and CM's respective internal decision rules translate preferences of individual members into the institutions' ideal points and, second, how these institutions interact in order to reach a joint decision.

Note that we here follow the—somewhat legalistic—game-theoretic analysis by Napel and Widgrén (2006), which argued that the European Commission is a formally powerless player in the codecision game. The reason is that—at least when transaction costs are zero—EP and CM can jointly enact *any* policy on which they agree in the so-called Conciliation Committee, without scope for a Commission veto. This implies that we can disregard the preferences of the Commission in the analysis. We will denote the ideal point that characterizes aggregate preferences of EP by π and that of CM by μ. Both are determined for a given issue by the respective pivotal player inside these decision-making bodies. The ordered individual ideal points of the members of the Council of Ministers will be denoted by $\mu_{(1)} \leq \ldots \leq \mu_{(28)}$; those of individual members of EP by $\pi_{(1)} \leq \ldots \leq \pi_{(766)}$.

2.2 Intra-Council Decision Making

The weighted voting system which is used for decision making in the Council was practically unchanged from the Treaty of Rome in 1957 until the Treaty of Nice in 2001. The Nice rules came into force on November 1, 2004 (at first in a somewhat modified transitional form), and basically maintained the old qualified majority voting (QMV) framework. However, it added two extra criteria, the so-called *safety-nets*, concerning the number of 'yes'-votes and the share of the total EU population which they represent. Specifically, the EU28 QMV requirement consists of three criteria: 260 out of 352 votes (73.9 %), a simple majority of member states (15 out of 28) and 62 % of the total EU population. The second and third requirements only have a negligible effect on possible winning coalitions (see, e.g., Baldwin et al. 2001, or Felsenthal and Machover 2001), and affect the quantitative results presented in Sect. 3 only at the 5th or 6th decimal place. The Nice voting weights are presented in Table 1 below.

The Lisbon Treaty's major revision to intra-Council voting rules is the switch from weighted voting to a dual majority system with an additional requirement for

[3] We extended the number of members from currently 754 to 766, based on the assumption that Croatia's 12 observers to EP will become actual members when Croatia joins the EU in July 2013.

Table 1 2012 population data, Nice weights, and power in EU28 under the Nice and Lisbon Treaty rules evaluated by the Shapley-Shubik and Penrose-Banzhaf indices

Member state	Population in 1,000s	Nice weight	SSI (Nice)	SSI (Lisbon)	PBI (Nice)	PBI (Lisbon)
Germany	81843.8	29	0.08580	0.14641	0.03039	0.17062
France	65397.9	29	0.08541	0.11089	0.03039	0.13826
United Kingdom	62989.6	29	0.08523	0.10616	0.03039	0.13380
Italy	60850.8	29	0.08518	0.10215	0.03039	0.13001
Spain	46196.3	27	0.07890	0.07542	0.02901	0.10259
Poland	38208.6	27	0.07841	0.06379	0.02901	0.08548
Romania	21355.8	14	0.03915	0.03969	0.01671	0.06506
Netherlands	16730.3	13	0.03613	0.03236	0.01562	0.05703
Greece	11290.8	12	0.03325	0.02418	0.01444	0.04786
Belgium	11041.3	12	0.03325	0.02379	0.01444	0.04744
Portugal	10541.8	12	0.03325	0.02306	0.01444	0.04660
Czech Republic	10504.2	12	0.03325	0.02301	0.01444	0.04653
Hungary	9962.0	12	0.03325	0.02222	0.01444	0.04562
Sweden	9482.9	10	0.02748	0.02154	0.01214	0.04482
Austria	8443.0	10	0.02748	0.02002	0.01214	0.04307
Bulgaria	7327.2	10	0.02748	0.01841	0.01214	0.04120
Denmark	5580.5	7	0.01908	0.01590	0.00857	0.03825
Slovakia	5404.3	7	0.01908	0.01567	0.00857	0.03795
Finland	5401.3	7	0.01908	0.01566	0.00857	0.03794
Ireland	4495.4	7	0.01908	0.01438	0.00857	0.03642
Croatia	4412.1	7	0.01908	0.01426	0.00857	0.03628
Lithuania	3199.8	7	0.01907	0.01254	0.00857	0.03422
Slovenia	2055.5	4	0.01094	0.01093	0.00491	0.03229
Latvia	2042.4	4	0.01094	0.01091	0.00491	0.03226
Estonia	1339.7	4	0.01092	0.00994	0.00491	0.03107
Cyprus	862.0	4	0.01092	0.00927	0.00491	0.03026
Luxembourg	524.9	4	0.01092	0.00879	0.00491	0.02969
Malta	420.1	3	0.00800	0.00865	0.00370	0.02951

blocking coalitions. A winning coalition must represent a majority of at least 55 % of EU member states and of 65 % of the total EU population. Moreover, the Lisbon Treaty prescribes that ‘no’ votes of at least four countries are needed in order to block proposals. However, the effect of this blocking clause is very small in any power computations.

Note that despite the fact that the Lisbon Treaty has already been in force since 1 December 2009, the new decision rule is not applied until November 2014. Furthermore, there will be a transition period from 1 November 2014 until 31 March 2017 during which any country can request the use of the old Nice Treaty rules. It seems plausible to assume that any country that is part of a coalition which is unsuccessful under the Lisbon rules (failing either to block or to pass legislation) but would be successful under the Nice rules will demand the use of the latter. So the former will probably not be used until April 2017.

As a benchmark for our analysis below, Table 1 also contains the intra-Council distribution of power in EU28 under the Nice and Lisbon Treaty voting rules according to the *Shapley-Shubik power index (SSI)* (Shapley 1953; Shapley and Shubik 1954). This index is closest in spirit to the strategic analysis pursued below, and in particular closer than the other main power index, the *Penrose-Banzhaf power index* (Penrose 1946; Banzhaf 1965). We will focus on the SSI and thus give traditional power analysis its 'best shot' in the later comparisons.[4] Many qualitative observations, e.g., that the Lisbon rules make the biggest four countries, and Romania and Malta more powerful than they were under the Nice rules, actually do not depend on which relative power measure is used.

2.3 The Codecision Procedure

The European Union's *codecision procedure* was introduced by the Maastricht Treaty, and initially applied to only 15 areas of Community activity. Its current version came into force in May 1999, introduced by the Treaty of Amsterdam. Its scope was already increased considerably under the Treaty of Amsterdam (May 1999) and the Treaty of Nice (February 2003). Under the Treaty of Lisbon, the codecision procedure entered a new era. It is now officially called the EU's *ordinary legislative procedure*. Moreover, its domain of application has been increased from 44 areas of Community activity under the Nice Treaty to 85 under the Lisbon Treaty. It applies to all previously covered areas such as environment, employment, social policies, education and consumer protection and to a number of new important areas such as agriculture, freedom, security, justice, common commercial policy and intellectual property.[5] The procedure is illustrated in Fig. 1. It is laid down in Art. 294 of the *Treaty on the Functioning of the European Union (TFEU)* and involves up to three readings of proposed legislation by EP and CM. It is initiated by a proposal of the European Commission, who can, however, be prompted by CM or EP to 'open the gates' (see Art. 225 and Art. 241 TFEU).[6] First, EP can approve this proposal or replace it with an amended version of its own. Then, CM either approves the proposal on the table or initiates a second stage of decision making by making amendments. This new proposal is either approved by EP or, again, amended. If in the latter case CM does not accept EP's proposal,[7] the *Conciliation Committee* represents a final chance to replace the status quo by a

[4] Corresponding values of the (absolute) Penrose-Banzhaf power index (PBI) are also included in Table 1 but are not explored in our comparative analysis.

[5] In some of these areas, EP had previously no say or only a right of consultation.

[6] Under the Treaty of Lisbon, proposals can—at least in specific cases—also be submitted on the initiative of a group of member states, on a recommendation by the European Central Bank, or at the request of the Court of Justice (see Art. 294(15) TFEU).

[7] The Commission—by a negative opinion on EP's proposal—can require CM to accept unanimously (see Art. 294(9) TFEU).

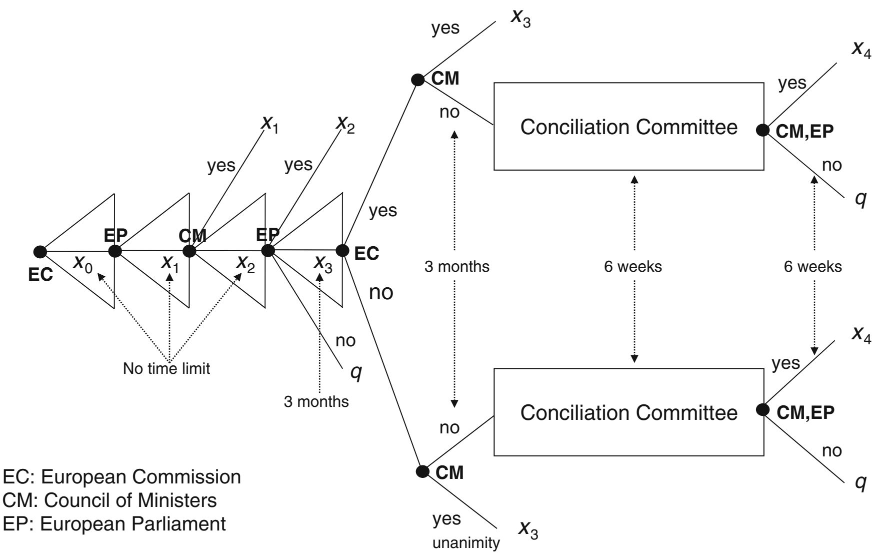


Fig. 1 Stylized codecision game tree (*Source* Napel and Widgrén 2006)

policy to its left or right. The Committee is composed of all 28 members of CM and an equally sized delegation of members of EP (MEPs). The committee is co-chaired by an EP Vice-President and the minister holding the Council Presidency without any fixed negotiation protocol. The Commission's formal role in the committee is only to facilitate agreement and to draft proposals requested by CM and EP. This is the key reason why the game-theoretic analysis of Napel and Widgrén (2006) did not consider it as a formally powerful player. If the CM and EP delegates agree on a compromise, it is submitted to CM and EP for acceptance in a third reading, in which both institutions use their standard qualified and simple majority voting rules, respectively. Note that the codecision procedure as such has not been changed by the Treaty of Lisbon. The only difference is that CM now decides by qualified majority in all policy domains that are covered by codecision. Under the Nice Treaty, there were some areas of Community activity for which CM had to decide by unanimity.

The bargaining outcome that EP and CM expect to result from invoking the Conciliation Committee plays a crucial strategic role at earlier stages of the procedure. Assuming complete information about preferences and using backward induction, one can conclude that it is indeed *the* determinant of any codecision agreement if all agents are strategic.

Accepted new legislation will usually come into effect at some date in the medium-term future. It is therefore reasonable to assume that neither EP nor CM has a pronounced preference for agreeing on a policy change a few weeks sooner rather than later. The codecision outcome can then be identified with the policy which CM and EP expect to agree on in Conciliation (either a new policy or the status quo). Therefore, quantitative analysis of EP's and CM's influence on codecision outcomes can be confined to the Conciliation stage.

We see neither empirical nor theoretical reasons to consider either EP or CM a more impatient or skilled bargainer. So we will use the *symmetric Nash bargaining solution* to predict the Conciliation agreement and thus, using backward induction, the codecision outcome. For our unidimensional policy space $X = [0, 1]$ and the benchmark case of utility that linearly decreases with distance, the symmetric Nash bargain corresponds to agreement on the institutional aggregate ideal point which is closer to the status quo whenever there are gains from trade, i.e., if both EP and CM want to move away from the status quo in the same direction (see Napel and Widgrén 2006 for a detailed derivation). Formally, we have

$$sign(q-\pi) = sign(q-\mu) \implies x^*(\pi,\mu) = \begin{cases} \pi; & d(\pi,q) \leq d(\mu,q) \\ \mu; & d(\pi,q) > d(\mu,q). \end{cases} \tag{1}$$

The Council's preferences, captured by its ideal point μ, are determined internally according to the Nice or Lisbon voting rules, which we discussed above. Looking at the Nice rules, let $w(\mu_{(i)})$ denote the number of votes (i.e., the voting weight) of the minister who has ideal point $\mu_{(i)}$, and $p(\mu_{(i)})$ the size of the population that he represents.[8] If CM considers a replacement of the status quo q by a policy to its left, the countries holding the left-most positions $\mu_{(1)}$, $\mu_{(2)}$, etc. will be the most enthusiastic about this. The critical CM member is the country that first brings about the required qualified majority as less and less enthusiastic supporters of a change are added to the coalition which endorses the new policy. We refer to this critical member as CM's *right pivot* R, and to its ideal point as CM's *right pivot position* μ_R. Under the Nice voting rules, the right pivot can be written as

$$R^{Nice} = \min\left\{r \in \{15,\ldots,28\} : \sum_{i=1}^{r} w(\mu_{(i)}) \geq 260 \wedge \sum_{i=1}^{r} p(\mu_{(i)}) \geq 0.62P^{EU}\right\}, \tag{2}$$

where P^{EU} refers to the EU's total population; we denote its ideal policy by $\mu_R^{Nice} \equiv \mu_{(R^{Nice})}$. This bliss point—reflecting the position of the government that is critical inside CM when coalition formation starts from the left-most position—is taken to be CM's aggregate position if the interaction with the European Parliament concerns a replacement of q by a policy to its left. It is the policy alternative that internally beats the status quo if that is sufficiently far to the right, and also beats any other status quo-beating policy.

[8] This already uses the fact that if countries' ideal points result from independent draws from a continuous probability distribution on X, such as the uniform one considered below, then there is almost surely only a single country with position $\mu_{(i)}$. So $w(\mu_{(i)})$ and $p(\mu_{(i)})$ will be well-defined with probability one.

Similarly, we have

$$L^{Nice} = \max\Big\{l \in \{1,\ldots,14\} : \sum_{i=l}^{28} w(\mu_{(i)}) \geq 260 \wedge \sum_{i=l}^{28} p(\mu_{(i)}) \geq 0.62P^{EU}\Big\}, \tag{3}$$

and $\mu_L^{Nice} \equiv \mu_{(L^{Nice})}$, reflecting the position of the government that is critical inside CM when coalition formation starts from the right-most position. It will be CM's aggregate position when a replacement of q by a policy to its right is contemplated.

Analogously, the Lisbon Treaty's voting rules lead to μ_R^{Lisbon} and μ_L^{Lisbon}, defined by

$$R^{Lisbon} = \min\Big\{\min\big\{r \in \{16,\ldots,28\} : \sum_{i=1}^{r} p(\mu_{(i)}) \geq 0.65P^{EU}\big\}, 25\Big\} \tag{4}$$

and

$$L^{Lisbon} = \max\Big\{\max\big\{l \in \{1,\ldots,13\} : \sum_{i=l}^{28} p(\mu_{(i)}) \geq 0.65P^{EU}\big\}, 4\Big\}. \tag{5}$$

They are the position variables which need to be considered regarding whether the double majority requirement inside CM (incl. the 4-blockers clause), which is prescribed by the Lisbon Treaty, is satisfied by some policy alternative. Note that, under either Treaty, *no* policy $x \in X$ would be supported by the required majority in CM if $\mu_L^j < q < \mu_R^j$.

Concerning the European Parliament, its 766 members need to approve any Conciliation compromise by simple majority. Entering negotiations with CM about some policy to the right of the status quo q, most of the potential positions of the EP delegation are such that a majority of MEPs would find it beneficial to intervene and select a different delegation. More concretely, consider the ordered MEP ideal points $\pi_{(1)} \leq \pi_{(2)} \leq \cdots \leq \pi_{(766)}$ and a provisional bargaining position π with $q < \pi < \pi_{(383)}$. Parliamentarians with ideal points $\pi_{(383)},\ldots,\pi_{(766)}$ then have the necessary majority and common interest to instead select some delegation with $\pi \geq \pi_{(383)}$ as EP's position for Conciliation negotiations. Similarly, MEPs with ideal points $\pi_{(1)},\ldots,\pi_{(384)}$ would block a position $\pi > \pi_{(384)}$. One can hence restrict EP's ideal point in negotiations about policies $x > q$ to $\pi \in [\pi_{(383)}, \pi_{(384)}]$. Recall that according to the Nash bargaining solution, it is the institution whose ideal point is closer to the status quo which is determining the Conciliation agreement. With this, and anecdotal evidence on EP's interest in being perceived as a powerful institution in the EU, in mind, we take the influence-maximizing ideal point $\pi = \pi_{(383)}$ to be EP's position in negotiations about $x > q$ and refer to the corresponding MEP as EP's *pivotal player*. By analogous reasoning, we identify EP with position $\pi = \pi_{(384)}$ for policies $x < q$.[9]

Note that, in principle, the internal position of EP need not coincide with the position taken by its delegation to the Conciliation Committee. In general, there could be gains from strategically picking a delegation whose interests diverge from the pivotal voter's (see, e.g., Segendorff 1998). However, under the above assumptions this cannot be advantageous by Eq. (1), any Conciliation agreement replacing the status quo amounts to the ideal point of either EP's or CM's delegation. Picking an EP delegation with a position to the left or right of its 'true' ideal point π thus has either no effect (CM's position is closer to status quo) or actually hurts EP's pivot. Namely, it may induce agreement on the distorted position π' instead of π when this would have been the outcome in the unmanipulated case, or it prevents agreement on the position of CM when that is actually closer to π than π' and hence preferable by EP's pivot.

It can be checked that negotiations in the Conciliation Committee can, for given preferences of MEPs and members of CM, never be simultaneously about policies $x > q$ *and* policies $x' < q$: if both institutions support, say, moving to the right of the status quo, i.e., both $\pi_{(383)}$ and $\mu_{(L)}$ lie to the right of q, then there is necessarily insufficient support for any $x < q$ because $\pi_{(384)} \geq \pi_{(383)}$ and $\mu_{(R)} \geq \mu_{(L)}$ must also lie to the right of q. This allows us to take μ as the well-defined ideal point of CM regarding any issue for which EP and CM want to move away from the status quo in the same direction, i.e., whenever both have an interest in reaching a deal. Note that μ is a function of countries' individual unordered ideal points $\mu_1, \ldots, \mu_{28}$—which give rise to the ordered ideal points $\mu_{(1)} \leq \cdots \leq \mu_{(28)}$—and the voting rule (either Nice or Lisbon).

3 Power Analysis

In order to obtain quantitative statements regarding the expected influence of individual Council members or the EP on EU decisions, we apply the framework proposed by Napel and Widgrén (2004) for the analysis of power in collective decision making.[10] It defines a player's *a priori power* in a given decision procedure and for a given probabilistic distribution of all relevant players' preferences as the *expected change to the equilibrium collective decision which would be brought about by a change in this player's preferences*. Alternatively, one could also make probabilistic assumptions about players' *actions*, rather than preferences

[9] Quite often in the spatial voting literature, EP is treated as a unitary actor. However, this simplification is not needed for the purposes of this paper. See Napel and Widgrén (2006) for robustness checks regarding the modeling of EP.

[10] That framework builds on—and very considerably generalizes—ideas which were first put forward by Widgrén and Napel (2002) (reprinted in this volume).

which induce actions. Traditional power indices take this 'short-cut' but thus lose the ability to transparently account for strategic interaction.

The framework links power analysis to the question: which impact would a marginal shift of a given player's ideal policy (caused, e.g., by a lobbyist, who evaluates all players' power before targeting any particular one) have on the collective decision? This approach to power measurement via a *sensitivity analysis* of collective decisions generalizes the weighted counting of players' pivot positions which is the basis of conventional power indices.

Before one can make statements about a priori power, one first needs to explicate and evaluate a posteriori power for a given preference profile. We will do so by considering the effect of a *marginal* shift of ideal points π or $\mu_1, \ldots, \mu_{28}$ to the left or right on the anticipated policy outcome. This effect is captured by the (partial) derivatives of the predicted outcome shown in Eq. (1) above. So the *a posteriori power* of EP, i.e., that for a *given* realization of status quo q and ideal points $\pi_1, \ldots, \pi_{766}$ and $\mu_1, \ldots, \mu_{28}$, is

$$\frac{\partial x^*(\pi, \mu, q)}{\partial \pi} = \begin{cases} 1 & \text{if } q < \pi < \mu \text{ or } \mu < \pi < q, \\ 0 & \text{otherwise}. \end{cases} \tag{6}$$

This formalizes that any (small) change of the player's ideal point with smaller status quo distance translates into a same-size shift of the agreed policy, provided there is agreement about changing the status quo at all.

What we are really interested in, however, is the *a priori power* of actors such as EP, namely, the influence not on a single issue but on average for many issues or—taking the 'veil of ignorance'—perspective of constitutional design—*in expectation*. In particular, our *strategic measure of power (SMP)*, derived from Napel and Widgrén (2004), is the expected impact that any marginal shift of EP's ideal policy π would have on the codecision outcome,

$$\xi_\pi = \Pr(\tilde{q} < \tilde{\pi} < \tilde{\mu}) + \Pr(\tilde{\mu} < \tilde{\pi} < \tilde{q}), \tag{7}$$

where $\tilde{q}$, $\tilde{\pi}$, and $\tilde{\mu}$ denote the random variables corresponding to status quo and institutional ideal points, and where a plausible a priori probability distribution of these random variables is assumed. For constitutional analysis, it is in our view most natural to assume that the individual ideal points of MEPs as well as of Council members are independently, identically, and—in line with the *principle of insufficient reason*, which is also invoked regarding player orderings or 'for'-or-'against' preferences by the Shapley-Shubik and Penrose-Banzhaf indices—*uniformly* distributed on the policy space X (here, the unit interval $[0, 1]$).[11] All our computations will hence be based on the a priori assumption of independent and uniformly distributed individual ideal points, as well as an independent and

[11] Assuming that all ideal points are mutually independent and uniformly distributed on $X = [0, 1]$ implies that $\tilde{\pi}$ is beta-distributed with parameters 383 and 384. The distribution of $\tilde{\mu}$ is considerably more complicated because of weighted voting.

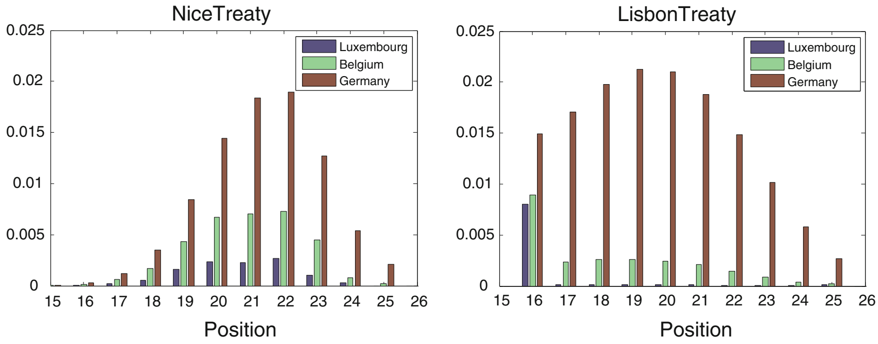


Fig. 2 Probability of being pivotal at a given position

uniformly distributed status quo q. Numerical results on EP's SMP, ξ_π, will be reported in the next section.

For an individual member k of CM, we obtain

$$\frac{\partial x^*(\pi, \mu(\mu_1, \ldots, \mu_{28}), q)}{\partial \mu_k} = \begin{cases} 1 & \textit{if } (q < \mu < \pi \textit{ or } \pi < \mu < q) \textit{ and } \mu = \mu_k, \\ 0 & \textit{otherwise} \end{cases} \tag{8}$$

as k's a posteriori power on a given issue, and

$$\xi_{\mu_k} = \left[\Pr(\tilde{q} < \tilde{\mu} < \tilde{\pi} \mid \tilde{\mu} = \tilde{\mu}_k) + \Pr(\tilde{\pi} < \tilde{\mu} < \tilde{q} \mid \tilde{\mu} = \tilde{\mu}_k)\right] \cdot \Pr(\tilde{\mu} = \tilde{\mu}_k) \tag{9}$$

as k's SMP, averaging over a large number of issues with independent and [0,1]-uniformly distributed individual ideal points and status quo.

Asymmetric voting weights in CM imply that conditioning on distinct events $\{\tilde{\mu} = \tilde{\mu}_i\}$ indeed affects the probability of event $\{\tilde{q} < \tilde{\mu} < \tilde{\pi}\}$. In particular, large countries with high voting weight are pivotal relatively more often in coalitions that already include many others—who have, on average, smaller weight—and, therefore, their associated spatial positions when being pivotal tend to be located more towards the extremes (e.g., quite far to the right if coalition-formation starts from the left). Figure 2 shows the probabilities

$$\Pr\left(\tilde{R}^j = i \ \wedge \ (\tilde{R}^j) = k\right)$$

of exemplary large, medium-sized, and small countries k being pivotal under the voting rules $j \in \{Nice, Lisbon\}$ at a particular rank position i when coalition-formation starts from the left, i.e., the chances to bring about the required qualified majority as the i-th member of a coalition that already includes $i-1$ members (with decreasing enthusiasm about changing the status quo towards the left).

Note that the Lisbon Treaty visibly shifts these distributions of pivotal positions to the left. More swings, thus, take place in smaller coalitions. This holds especially for smaller countries. The explanation is two-fold. First, for small countries the membership criterion is a much more important source of influence than the population criterion. That explains why Belgium and Luxembourg have most of

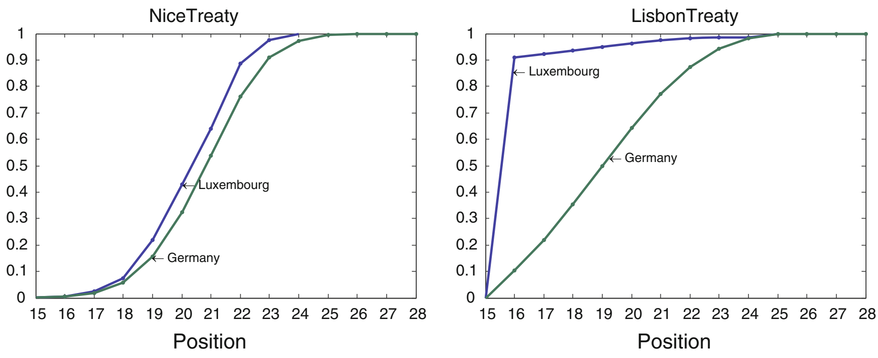


Fig. 3 Conditional cumulative probability of being pivotal at a given position

their pivotal positions at $i = 16$, which is exactly the effective Lisbon membership threshold. Second, for big countries like Germany the population criterion contributes more to power. The criterion involves a threshold of slightly more than 315 million citizens. The rank distribution of Germany's pivotal positions has its mode at rank 19, which in expected terms corresponds to 345 millions (using the average population per country). With its population of nearly 82 millions, Germany is easily able to swing winning coalitions which pass the population threshold into losing ones.

Figure 3 illustrates the corresponding *conditional* pivotal position distributions (in cumulative terms) for a small and a large country, Luxembourg and Germany. Already under the Nice rules, Germany has its (a priori random) right pivot positions more to the right than Luxembourg; it is pivotal at a rank position that is larger than that of Luxembourg in the sense of first-order stochastic dominance. This becomes much more pronounced under the Lisbon rules, because their focus on population sizes makes Germany an even 'larger' player, relatively speaking, than it was under the Nice rules—meaning that a random coalition of fixed size $i - 1$ is typically farther away from passing the voting thresholds than under the Nice rules; so that a greater coalition size $i - 1$ is compatible with becoming a winning coalition only after Germany joins. In summary, small countries exert power in coalitions with relatively few members (which are relatively big countries), and big countries exert power in coalitions containing a relatively high number of small countries.

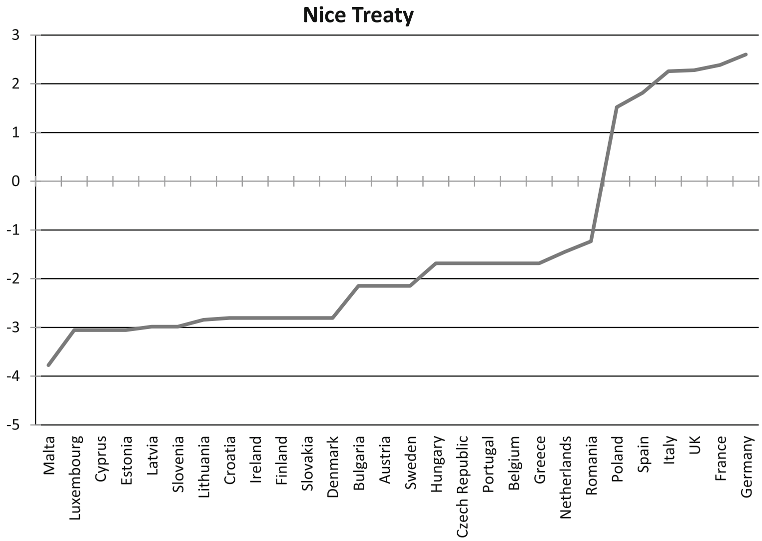
4 Results

Table 2 reports the SMP values of individual Council members and the EP for EU28. It also shows, as a measure of *relative* power inside CM, the *normalized SMP values (NSMP)*, and the relative differences between the SSI values (cf. Table 1) and the intra-CM power assessment implied by the NSMP.

Table 2 Strategic power in EU28 under Nice and Lisbon Treaty rules and the intra-CM difference to SSI in the codecision procedure (EP as 766 MEPs)

Member state	SMP (Nice)	SMP (Lisbon)	NSMP (Nice)	NSMP (Lisbon)	(SSI–NSMP)/ SSI % (Nice)	(SSI – NSMP)/ SSI % (Lisbon)
Germany	0.04416	0.08283	0.08357	0.14522	2.60192	0.81434
France	0.04405	0.06284	0.08337	0.11017	2.38701	0.64894
United Kingdom	0.04401	0.06024	0.08329	0.10561	2.27898	0.51831
Italy	0.04400	0.05801	0.08326	0.10169	2.25532	0.45188
Spain	0.04094	0.04274	0.07747	0.07493	1.81234	0.64801
Poland	0.04080	0.03600	0.07722	0.06311	1.52140	1.05703
Romania	0.02094	0.02284	0.03963	0.04003	-1.23165	-0.86731
Netherlands	0.01937	0.01868	0.03665	0.03274	-1.44866	-1.19387
Greece	0.01786	0.01398	0.03381	0.02450	-1.68660	-1.31526
Belgium	0.01786	0.01375	0.03381	0.02411	-1.68660	-1.36007
Portugal	0.01786	0.01333	0.03381	0.02337	-1.68660	-1.35066
Czech Republic	0.01786	0.01330	0.03381	0.02332	-1.68660	-1.34791
Hungary	0.01786	0.01284	0.03381	0.02252	-1.68660	-1.32828
Sweden	0.01483	0.01245	0.02807	0.02182	-2.14917	-1.28365
Austria	0.01483	0.01157	0.02807	0.02028	-2.14917	-1.27598
Bulgaria	0.01483	0.01063	0.02807	0.01864	-2.14917	-1.25370
Denmark	0.01037	0.00917	0.01962	0.01608	-2.80732	-1.12676
Slovakia	0.01037	0.00903	0.01962	0.01583	-2.80732	-1.07545
Finland	0.01037	0.00903	0.01962	0.01583	-2.80732	-1.07456
Ireland	0.01037	0.00828	0.01962	0.01452	-2.80732	-0.97894
Croatia	0.01037	0.00821	0.01962	0.01440	-2.80732	-0.97046
Lithuania	0.01036	0.00721	0.01961	0.01264	-2.84373	-0.75601
Slovenia	0.00595	0.00627	0.01126	0.01099	-2.98311	-0.53928
Latvia	0.00595	0.00626	0.01126	0.01097	-2.98311	-0.53128
Estonia	0.00595	0.00569	0.01126	0.00997	-3.05496	-0.27216
Cyprus	0.00595	0.00529	0.01126	0.00928	-3.05496	-0.09668
Luxembourg	0.00595	0.00501	0.01126	0.00879	-3.05496	0.01110
Malta	0.00439	0.00493	0.00830	0.00864	-3.77756	0.07958
Council aggregate	0.52842	0.57042	1.00000	1.00000	0.00000	0.00000
European Parliament	0.01990	0.11542	n.a.	n.a.	n.a.	n.a.

First, it is an important observation that the relative differences between the SSI and NSMP values are not substantial. The maximal deviation is 3.8 % (for Malta, under the Nice rules). In general, the relative differences, either positive or negative, are bigger for small countries under the Nice rules and for medium-sized countries under the Lisbon rules. This may seem natural on the one hand since their SSI values are smaller; on the other hand this is mainly linked to the two types of situations which cause SSI and NSMP to diverge: some of a country's swing positions that enter the SSI are not counted by the SMP because either (a) EP rather CM is pivotal for a particular preference configuration, or (b) EP and CM cannot agree on a replacement of the status quo by a policy to its left or right.



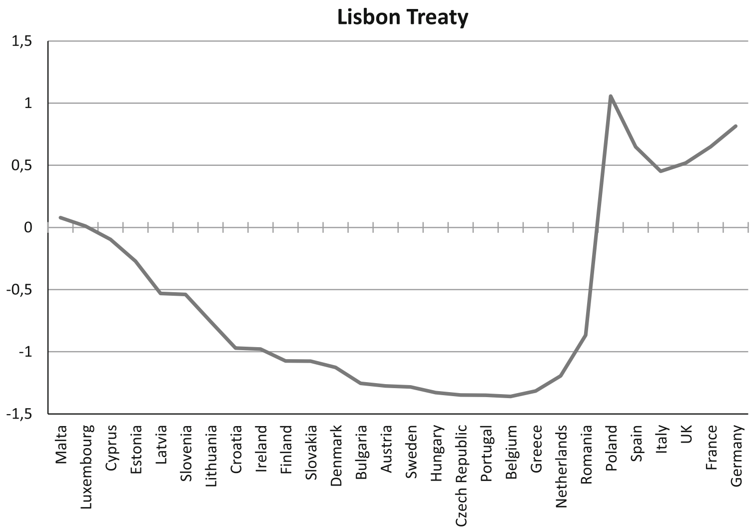


Fig. 4 Relative differences (%) between the SSI and NSMP under Nice and Lisbon rules

Second, compared to our earlier findings, the inter-institutional distribution of power between EP and CM is practically unaffected by taking the true voting weight distribution in CM into account. At the aggregate inter-institutional level, the effect of intra-institutional voting weights is very small (in contrast to the intra-institutional quota), as already argued in Napel and Widgrén (2006). In particular, the difference between 0.528 vs. 0.020 reported as CM's and EP's respective SMP value here, and 0.590 versus 0.023 reported in Napel and Widgrén (2006) is mainly due to the increase of CM's quota from about 72.2 % of total weight for EU25 to 73.9 % for EU28: a higher quota decreases the probability that a replacement of q can be agreed on by CM and EP; this reduces CM's expected influence on the outcome by more than EP.

Third, the patterns of deviations between SSI and NSMP under the Lisbon and Nice Treaties differ regarding their monotonicity properties. Under the Nice rules, the SSI underestimates the relative power of all countries that are smaller than Poland. The top panel of Fig. 4 shows monotonically increasing percentages.[12] The bottom panel illustrates non-monotonic deviations under the Lisbon rules: the SSI underestimates the relative power of most countries that are smaller than Poland but not that of the very smallest ones, Luxembourg and Malta.[13]

How the pattern of the relative differences between SSI and NSMP values depends on the voting rules relates to the two types of 'miscountings', (a) and (b), of pivot positions by the SSI described above. Under the Lisbon rules, small countries' pivotal positions are more concentrated in the relative middle of X, since they mostly matter due to the membership criterion (see Figs. 2 and 3 above). EP's position π is a priori highly concentrated in the middle of X, too. This means that it will be relatively often the case that EP's ideal point π is closer to the status quo q than is CM's (= the small country's) ideal point μ. Many of a small country's pivot positions hence do not translate into actual influence on collective decisions. Because of their comparatively more central conditional pivot position distribution, small countries lose relatively more of their intra-CM swings in this way than do large countries, i.e., the effect of (a) is more pronounced for them than for large countries.

However, having a more extreme conditional distribution of one's pivot positions, as large countries do, also comes with a disadvantage, namely a greater probability that the random status quo q is situated between π and μ. That means that EP and CM cannot agree on a replacement of the status quo, and that the corresponding purely intra-CM pivot position does not translate into influence on the collective decision either. Large countries lose relatively more of their intra-CM swing positions, which are counted by the SSI, in this way, i.e., the effect of (b) is more pronounced for them than for small countries. Whether effect (a) or (b) dominates, and hence whether small or large countries' relative power is overstated by the SSI, depends on *how much* more to the extremes of X the pivot positions of large countries are located on average.

[12] Note that underestimations require corresponding overestimations by the definition of *relative* power.

[13] The pattern of the Lisbon deviations is qualitatively different from its EU27 analogue in Napel and Widgrén (2009). The explanation seems to be that the Lisbon rule's 55 %-requirement amounted to an effective member quota of about 55.56 % for EU27 (15 out of 27 members), while the latter is around 57.14 % for EU28 (16 out of 28 members). This implicit quota rise somewhat increases the effect of (b), reduces the net effect of (a) and (b), and renders the Lisbon deviations more similar to the Nice deviations for EU28.

The higher decision quota under the Nice rules makes it more likely that EP and CM do not see mutual gains from striking a deal than under the Lisbon rules; so the magnitude of effect (b) relative to (a) is greater under the Nice Treaty. It turns out that (b) dominates (a) under the Nice rules, i.e., large CM members lose a greater share of their swings due to strategic interaction with EP than small members; hence the SSI understates the latter's true strategic power. For the Lisbon rules, the lower quota reduced the importance of (b) sufficiently to let (a) become dominant in EU27. But this is no longer the case for EU28 (see fn. 13). So the resulting signs of the deviations are more similar to the Nice rules for EU28. The remaining differences are linked to the fact that pivotality is driven essentially only by the weight criterion under the Nice rules, while it is driven by both the weight and membership criteria for the Lisbon rules.

5 Concluding Remarks

In this paper, we have studied the intra-institutional distribution of power in the Council of Ministers assuming spatial preferences and strategic interaction with the EP according to the EU's codecision or ordinary legislative procedure for EU28. We have first derived the equilibrium outcome of the procedure considering arbitrary but fixed spatial preferences, and have then randomized these preferences in order to conduct a priori power analysis with constitutional relevance. To our knowledge, this study is the first to consider weighted voting in a power analysis of EU28 which is based on a procedural voting model.

The paper allows three main conclusions: first, the numerical differences to our earlier assessment of the power relation between Council and Parliament turn out to be quite small. Disregarding intra-institutional voting weights in the study of inter-institutional power relations delivers a pretty good first approximation. As already found in Napel and Widgrén (2006), the distribution of power between EP and CM is very uneven: CM is by an order of magnitude more influential on codecision outcomes than EP a priori. From an a priori perspective, big individual member states like Germany are under the Nice rules more powerful than the European Parliament. The Lisbon Treaty rules, however, improve EP's power position considerably.

Secondly, the relative differences between the standard Shapley-Shubik index and our normalized measure of strategic power are not very big. This should, however, not be an excuse for continuing to disregard procedures and strategies in future analysis of intra-institutional power. There are clear differences between strategic power across different procedures which are not picked up by standard indices. For example, the individual power differences between the codecision procedure and the EU's *consultation procedure* are large; they significantly exceed the differences between the Nice vs. Lisbon intra-CM voting rule results reported here. Classical power measures should only serve as a first approximation. Of course, they can also be applied safely to situations where there is *no* inter-institutional interaction, or where agenda setting and amendment procedures are irrelevant. But neither is the case for the European Union.

Finally, as a methodological corollary to the second conclusion, we find that the criticism of Garrett and Tsebelis has its justification, but proves to be far less important numerically than conceptually. While our strategic measure of power is, figuratively, a quite distant cousin of the SSI—and corrects certain problems of traditional indices by explicitly modeling procedures and strategic interaction—it turns out that, at least in the context of the EU's codecision procedure, SMP and SSI are very close.[14]

References

Baldwin, R., & Widgrén, M. (2004). Winners and losers under various dual majority rules for the EU Council of Ministers. In M. Wiberg (Ed.), *Reasoned choices—Essays in honor of academy Professor Hannu Nurmi on the occasion of his 60th birthday*. Turku: The Finnish Political Science Association.

Baldwin, R. E., Berglöf, E., Giavazzi, F., & Widgrén, M. (2001). *Nice try: Should the treaty of Nice be ratified? Monitoring european integration 11*. London: Center for Economic Policy Research.

Banzhaf, J. F. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review*, *19*(2), 317–343.

Felsenthal, D., & Machover, M. (2001). The treaty of Nice and qualified majority voting. *Social Choice and Welfare*, *18*(3), 431–464.

Felsenthal, D., & Machover, M. (2004). Analysis of QM rules in the draft Constitution for Europe proposed by the European Convention, 2003. *Social Choice and Welfare*, *23*(1), 1–20.

Garrett, G., & Tsebelis, G. (1999). Why resist the temptation to apply power indices to the European Union? *Journal of Theoretical Politics*, *11*(3), 291–308.

Laruelle, A., & Widgrén, M. (1998). Is the allocation of power among EU states fair? *Public Choice*, *94*(3–4), 317–340.

Leech, D. (2002). Designing the voting system for the EU Council of Ministers. *Public Choice*, *113*(3–4), 437–464.

Napel, S., & Widgrén, M. (2004). Power measurement as sensitivity analysis—a unified approach. *Journal of Theoretical Politics*, *16*(4), 517–538.

Napel, S., & Widgrén, M. (2006). The inter-institutional distribution of power in EU codecision. *Social Choice and Welfare*, *27*(1), 129–154.

Napel, S., & Widgrén, M. (2008). Shapley-Shubik versus strategic power—live from the UN Security Council. In M. Braham & F. Steffen, (Eds.), *Power, freedom and voting—essays in honor of Manfred J. Holler* (pp. 99–118). Berlin: Springer-Verlag.

Napel, S., & Widgrén, M. (2009). Strategic power in the European Union's codecision procedure. *Homo Oeconomicus 26*(3–4), 297–316. Essays in honor of Hannu Nurmi, vol. 1.

Napel, S., & Widgrén, M. (2011). Strategic vs. non-strategic voting power in the EU Council of Ministers: The consultation procedure. *Social Choice and Welfare*, *37*(3), 511–541.

Penrose, L. S. (1946). The elementary statistics of majority voting. *Journal of the Royal Statistical Society*, *109*(1), 53–57.

[14] See Napel and Widgrén (2008, 2011) for a broadly similar conclusion derived from strategic analysis of a priori power in the UN Security Council and the EU's consultation procedure.

Segendorff, B. (1998). Delegation and threat in bargaining. *Games and Economic Behavior*, *23*, 266–283.
Shapley, L. S. (1953). A value for *n*-person games. In H. W. Kuhn & A. W. Tucker (Eds.), *Contributions to the theory of games* (Vol. II, pp. 307–317). Princeton, NJ: Princeton University Press.
Shapley, L. S., & Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review*, *48*(3), 787–792.
Steunenberg, B. (1994). Regulatory policymaking in a parliamentary setting. In P. Herder-Dorneich, K.-H. Schenk & D. Schmidtchen (Eds.), *Jahrbuch für Neue Politische Ökonomie* (Vol. 13, pp. 36–57). Tübingen: Mohr-Siebeck.
Steunenberg, B., & Selck, T. (2006). Testing procedural models of EU legislative decision-making. In R. Thomson, F. Stokman, C. Achen, & T. König (Eds.), *The EU decides* (pp. 54–85). Cambridge: Cambridge University Press.
Tsebelis, G. (1994). The power of the European Parliament as a conditional agenda setter. *American Political Science Review*, *88*(1), 128–142.
Widgrén, M. (1994). Voting power in the EC and the consequences of two different enlargements. *European Economic Review*, *38*(5), 1153–1170.
Widgrén, M., & Napel, S. (2002). The power of a spatially inferior player. *Homo Oeconomicus*, *19*(3), 327–343.

Square Root Voting System, Optimal Threshold and π

Karol Życzkowski and Wojciech Słomczyński

1 Introduction

Recent political debate on the voting system used in the Council of Ministers of the European Union stimulated research in the theory of indirect voting, see e.g. (Felsenthal and Machover 2001; Leech 2002; Andjiga et al. 2003; Pajala and Widgrén 2004; Beisbart et al. 2005). The *double majority* voting system, adopted for the Council by The Treaty of Lisbon in December 2007 is based on two criteria: 'per capita' and 'per state'. This system apparently reflects the principles of equality of Member States and that of equality of citizens. However, as recently analyzed by various authors Baldwin and Widgrén (2004), Ade (2003), Słomczyński and Życzkowski (2006), Algaba et al. (2007), Hosli (2008), Bârsan-Pipu and Tache (2009), Kirsch (2010), Moberg (2010), Leech and Aziz (2010), Pukelsheim (2010), Słomczyński and Życzkowski (2010), in such a system the large states gain a lot of power from the direct link to population, while the smallest states derive disproportionate power from the other criterion.

This chapter was presented to the *Voting Power in Practice Symposium* at the London School of Economics, 20–22 March 2011, sponsored by the Leverhulme Trust.

K. Życzkowski (✉)
Institute of Physics, Jagiellonian University, ul.Reymonta 4 30-059 Kraków, Poland
e-mail: karol@tatry.if.uj.edu.pl

Center for Theoretical Physics, Polish Academy of Sciences, Al. Lotników 32/46 02-668 Warszawa, Poland

W. Słomczyński
Institute of Mathematics, Jagiellonian University, ul. Lojasiewicza 6, Kraków 30-348, Poland
e-mail: wojciech.slomczynski@im.uj.edu.pl

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_30,

The combined effect saps influence away from all medium-sized countries. Ironically, a similar conclusion follows from a book by Lionel Penrose, who wrote already in 1952 (Penrose 1952):

If two votings were required for every decision, one on a per capita basis and the other upon the basis of a single vote for each country, this system would be inaccurate in that it would tend to favor large countries.

To quantify the notion of voting power, mathematicians introduced the concept of power index of a member of the voting body, which measures the probability that his vote will be decisive in a hypothetical ballot: Should this member decide to change his vote, the winning coalition would fail to satisfy the qualified majority condition. Without any further information about the voting body it is natural to assume that all potential coalitions are equally likely. This very assumption leads to the concept of Banzhaf(-Penrose) index called so after John Banzhaf, an American attorney, who introduced this index independently in 1965 (Banzhaf 1965).

Note that this approach is purely normative, not descriptive: we are interested in the potential voting power arising from the voting procedure itself. Calculation of the voting power based on the counting of majority coalitions is applicable while analyzing institutions in which alliances are not permanent, but change depending upon the nature of the matter under consideration.

To design a representative voting system, i.e. the system based on the democratic principle, that the vote of any citizen of any Member State is of equal worth, one needs to use a weighted voting system. Consider elections of the government in a state with population of size N. It is easy to imagine that an average German citizen has smaller influence on the election of his government than, for example, a citizen of the neighboring Luxembourg. Analyzing this problem in the context of voting in the United Nations just after the World War II Penrose showed, under some natural assumptions, that in such elections the voting power of a single citizen decays as one over square root of N. Thus, the system of indirect voting applied to the Council is representative, if the voting power of each country is proportional to the square root of N, so that both factors cancel out. This statement is known in the literature under the name of the *Penrose square root law* (Penrose 1946; Felsenthal and Machover 1998). It implies that the voting power of each member of the EU Council should behave as $\sqrt{N}$ and such voting systems have been analyzed in this context by several experts since late 1990s (Felsenthal and Machover 1997; Laruelle and Widgrén 1998).

It is challenging to explain this fact in a way accessible to a wide audience (Życzkowski et al. 2006; Kirsch et al. 2007; Pukelsheim 2007; Pöppe 2007). A slightly paradoxical nonlinearity in the result of Penrose is due to the fact that voting in the Council should be considered as a two-tier voting system: Each member state elects a government, which delegates its representative to the Council. Any representative has to say *'Yes'* or *'No'* on behalf of his state in every voting organized in the Council. The key point is that in such a voting each member of the Council cannot split his vote. Making an idealistic assumption that the vote of a Minster in the Council represents the will of the majority of the citizens of the state he represents, his vote '*Yes*' means only that a majority of the population of his state supports this decision, but does not reflect the presence of a minority.

Consider an exemplary issue to be voted in the Council and assume that the preferences of the voters in each state are known. Assume hypothetically that a majority of population of Malta says '*Yes*' on a certain issue, the votes in Italy split as 30 millions '*Yes*' and 29 millions '*No*', while all 43 millions of citizens of Spain say '*No*'. A member of the Council from Malta follows the will of the majority in his state and votes '*Yes*'. So does the representative of Italy. According to the double majority voting system his vote is counted on behalf of the total number of 59 millions of the population of Italy. Thus these voting rules allow 30 millions of voters in Italy to over-vote not only the minority of 29 millions in their state (which is fine), but also, with the help of less than half a million of people from Malta, to over-vote 43 millions of Spaniards.

This pedagogical example allows one to conclude that the double majority voting system would work perfectly, if all voters in each member state had the same opinion on every issue. Obviously such an assumption is not realistic, especially in the case of the European states, in which the citizens can nowadays afford the luxury of an independent point of view. In general, if a member of the Council votes '*Yes*' on a certain issue, in an ideal case one may assume that the number of the citizens of his state which support this decision varies from 50 to 100 % of the total population. In practice, no concrete numbers for each state are known, so to estimate the total number of European citizens supporting a given decision of the Council one has to rely on statistical reasoning.

To construct the voting system in the Council with voting powers proportional to the square root of populations one can consider the situation, where voting weights are proportional to the square root of populations and the Council takes its decision according to the principle of a *qualified majority*. In other words, the voting in the Council yields acceptance, if the sum of the voting weights of all Ministers voting '*Yes*' exceeds a fixed quota q, set for the qualified majority. From this perspective the quota q can be treated as a free parameter (Leech and Machover 2003; Machover 2010), which may be optimized in such a way that the mean discrepancy Δ between the voting power (measured by the Banzhaf index) and the voting weight of each member state is minimal.

In the case of the population in the EU consisting of 25 member states it was shown (Słomczyński and Życzkowski 2004; Życzkowski et al. 2006) that the value of the optimal quota q_* for qualified majority in the Penrose's square root system is equal to 62 %, while for EU-27 this number drops down to 61.5 % (Słomczyński and Życzkowski 2006, 2007). Furthermore, the optimal quota can be called *critical*, since in this case the mean discrepancy $\Delta(q_*)$ is very close to zero and thus the voting power of every citizen in each member state of the Union is practically equal. This simple scheme of voting in the EU Council based on the square root law of Penrose supplemented by a rule setting the optimal quota to q_* happens to give larger voting powers to the largest EU than the Treaty of Nice, but smaller ones than the Treaty of Lisbon. Therefore this voting system has been dubbed by the media as the *Jagiellonian Compromise*.

It is known that the existence of the critical quota q_*, is not restricted to the particular distribution of the population in the European Union, but it is also characteristic of a generic distribution of the population (Słomczyński and Życzkowski 2004; Chang et al. 2006; Słomczyński and Życzkowski 2006). The value of the critical quota depends on the particular distribution of the population in the 'union', but even more importantly, it varies considerably with the number M of member states. An explicit approximate formula for the critical quota was derived in Słomczyński and Życzkowski (2007). It is valid in the case of a relatively large number of the members of the 'union' and in the asymptotic limit, $M \to \infty$, the critical quota tends to 50 %, in consistence with the so-called *Penrose limit theorem* (Lindner and Machover 2004).

On one hand it is straightforward to apply this explicit formula for the current population of all member states of the existing European Union, as well as to take into account various possible scenarios of a possible extension of the Union. On the other hand, if the number of member states is fixed, while their populations vary in time, continuous update of the optimal value for the qualified majority may be cumbersome and unpractical. Hence one may try to neglect the dependence on the particular distribution of the population by selecting for the quota the mean value of $\langle q \rangle$, where the average is taken over a sample of random population distributions, distributed uniformly in the allowed space of M-point probability distributions. In this work we perform such a task and derive an explicit, though approximate, formula for the average critical quota.

This chapter is organized as follows. In Sect. 2 devoted to the one-tier voting system, we recall the definition of Banzhaf index and review the Penrose square root law. In Sect. 3, which concerns the two-tier voting systems, we describe the square root voting system and analyze the average number of misrepresented voters. Section 4 is devoted to the problem of finding the optimal quota for the qualified majority. It contains the key result of this paper: derivation of a simple approximate formula for the average optimal quota, which depends only on the number M of the member states and is obtained by averaging over an ensemble of random distributions of the population of the 'union'.

2 One Tier Voting

Consider a voting body consisting of M voters voting according to the qualified majority rule. Assume that the weights of the votes need not to be equal, which is typical e.g. in the case of an assembly of stockholders of a company: the weight of the vote of a stockholder depends on the number of shares he or she possesses. It is worth to stress that, generally, the voting weights do not directly give the voting power.

To quantify the a priori voting power of any member of a given voting body game theorists introduced the notion of a power index. It measures the probability that a member's vote will be decisive in a hypothetical ballot: should this player decide to change its vote, the winning coalition would fail to satisfy the qualified majority condition. In the game theory approach to voting such a player is called *pivotal*.

The assumption that all potential coalitions of voters are equally likely leads to the concept of the Banzhaf index (Penrose 1946; Banzhaf 1965). To compute this power index for a concrete case one needs to enumerate all possible coalitions, identify all winning coalitions, and for each player find the number of cases in which his vote is decisive.

Let M denote the number of voters and ω the total number of all winning coalitions, that satisfy the qualified majority condition. Assume that ω_k denotes the number of winning coalitions that include the kth player; where $k = 1, \ldots, M$. Then the Banzhaf index of the kth voter reads

$$\psi_k := \frac{\omega_k - (\omega - \omega_k)}{2^{M-1}} = \frac{2\omega_k - \omega}{2^{M-1}}. \tag{1}$$

To compare these indices for decision bodies consisting of different number of players, it is convenient to define the *normalized Banzhaf (-Penrose) index*:

$$\beta_k := \frac{\psi_k}{\sum_{i=1}^{M} \psi_i} \tag{2}$$

such that $\sum_{i=1}^{M} \beta_i = 1$.

In the case of a small voting body such a calculation is straightforward, while for a larger number of voters one has to use a suitable computer program.

2.1 Square Root Law of Penrose

Consider now the case of N members of the voting body, each given a single vote. Assume that the body votes according to the standard majority rule. On one hand, since the weights of each voter are equal, so must be their voting powers. On the other hand, we may ask, what happens if the size N of the voting body changes, for instance, if the number of eligible voters gets doubled, how does this fact influence the voting power of each voter?

For simplicity assume for a while that the number of voters is odd, $N = 2j + 1$. Following original arguments of Penrose we conclude that a given voter will be able to effectively influence the outcome of the voting only if the votes split half and half: If the vote of j players would be '*Yes*' while the remaining j players vote '*No*', the role of the voter we analyze will be decisive.

Basing upon the assumption that all coalitions are equally likely one can ask, how often such a case will occur? In mathematical language the model in which this assumption is satisfied is equivalent to the *Bernoulli scheme*. The probability that out of $2j$ independent trials we obtain k successes reads

$$P_k := \binom{2j}{k} p^k (1-p)^{2j-k}, \tag{3}$$

where p denotes the probability of success in each event. In the simplest symmetric case we set $p = 1 - p = 1/2$ and obtain

$$P_j \;=\; \left(\frac{1}{2}\right)^{2j} \frac{(2j)!}{(j!)^2}. \tag{4}$$

For large N we may use the Stirling approximation for the factorial and obtain the probability ψ that the vote of a given voter is decisive

$$\psi = P_j \;\sim\; 2^{-2j} \frac{(2j/e)^{2j}\sqrt{4\pi j}}{[(j/e)^j\sqrt{2\pi j}]^2} \;=\; \frac{1}{\sqrt{\pi j}} \;\sim\; \sqrt{\frac{2}{\pi N}}. \tag{5}$$

For N even we get the same approximation. In this way one can show that the voting power of any member of the voting body depends on its size as $1/\sqrt{N}$, which is the *Penrose square root law*. The above result is obtained under the assumption that the votes of all citizens are uncorrelated. A sound mathematical investigation of the influence of possible correlations between the voting behavior of individual citizens for their voting power has been recently presented by Kirsch (2007). It is easy to see that due to strong correlations certain deviations from the square root law have to occur, since in the limiting case of unanimous voting in each state (perfect correlations), the voting power of a single citizen from a state with population N will be inversely proportional to N.

The issue that the assumptions leading to the Penrose law are not exactly satisfied in reality was raised many times in the literature, see, e.g. (Gelman et al. 2002, 2004), also in the context of the voting in the Council of the European Union (Laruelle and Valenciano 2008). However, it seems not to be easy to design a rival model voting system which correctly takes into account the essential correlations, varying from case to case and evolving in time. Furthermore, it was argued (Kirsch 2007) that the strength of the correlations between the voters tend to decrease in time. Thus, if one is to design a voting system to be used in the future in the Council of the European Union, it is reasonable to consider the idealistic case of no correlations between individual voters. We will follow this strategy and in the sequel rely on the square root law of Penrose.

2.2 Pivotal Voter and the Return Probability in a Random Walk

It is worth to emphasize that the square root function appearing in the above derivation is typical to several other reasonings in mathematics, statistics and physics. For instance, in the analyzed case of a large voting body, the probability distribution P_k in the Bernoulli scheme can be approximated by the Gaussian distribution with the standard deviation being proportional to $1/\sqrt{N}$. It is also instructive to compare the above voting problem with a simple model of a random walk on the one dimensional lattice.

Assume that a particle subject to external influences in each step jumps a unit distance left or right with probability one half. What is the probability that it returns to the initial position after N steps? It is easy to see that the probability scales as $1/\sqrt{N}$, since the answer is provided by exactly the same reasoning as for the Penrose law.

Consider an ensemble of particles localized initially at the zero point and performing such a random walk on the lattice. If the position of a particle at time n differs from zero, in half of all cases it will jump towards zero, while in the remaining half of cases it will move in the opposite direction. Hence the *mean* distance $\langle D \rangle$ of the particle from zero will not change. On the other hand, if at time n the particle happened to return to the initial position, in the next step it would certainly jump away from it, so the mean distance from zero would increase by one.

To compute the mean distance form zero for an ensemble of random particles performing N steps, we need to sum over all the cases, when the particle returns to the initial point. Making use of the previous result, that the return probability $P(n)$ at time n behaves as $1/\sqrt{n}$, we infer that during the time N the mean distance behaves as

$$\langle D(N) \rangle \approx \sum_{n=1}^{N} P(n) \approx \sum_{n=1}^{N} \frac{1}{\sqrt{n}} \sim \sqrt{N}. \tag{6}$$

This is just one formulation of the *diffusion law*. As shown, the square root of Penrose is closely related with some well known results from mathematics and physics, including the Gaussian approximation of binomial distribution and the diffusion law.

3 Two Tier Voting

In a two-tier voting system each voter has the right to elect his representative, who votes on his behalf in the upper chamber. The key assumption is that, on one hand, he should represent the will of the population of his state as best he can, but, on the other hand, he is obliged to vote '*Yes*' or '*No*' in each ballot and cannot split his vote. This is just the case of voting in the Council of the EU, since citizens in each member state choose their government, which sends its Minister to represent the entire state in the Council.

These days one uses in the Council the triple majority system adopted in 2001 in the *Treaty of Nice*. The Treaty assigned to each state a certain number of 'weights', distributed in an ad hoc fashion. The decision of the Council is taken if the coalition voting in favour of it satisfies three conditions:

(a) it is formed by the standard majority of the member states;
(b) states forming the coalition represent more then 62 % of the entire population of the Union;
(c) the total number of weights of the '*Yes*' votes exceeds a quota equal to approximately 73.9 % of all weights.

Although all three requirements have to be fulfilled simultaneously, detailed analysis shows that condition (c) plays a decisive role in this case: if it is satisfied, the two others will be satisfied with a great likelihood as well (Felsenthal and Machover 2001; Leech 2002).

Therefore, the voting weights in the Nice system play a crucial role. However, the experts agree (Felsenthal and Machover 2001; Pajala and Widgrén 2004) that the choice of the weights adopted is far from being optimal. For instance the voting power of some states (including e.g. Germany and Romania) is significantly smaller than in the square root system. This observation is consistent with the fact that Germany was directly interested to abandon the Nice system and push toward another solution that would shift the balance of power in favor of the largest states.

In the double majority voting system, adopted in December 2007 in Lisbon, one gave up the voting weights used to specify the requirement (c) and decided to preserve the remaining two conditions with modified majority quotas. A coalition is winning if:

(a') it is formed by at least 55 % of the members states;
(b') it represents at least 65 % of the population of the Union;

Additionally, every coalition consisting of all but three (or less) countries is winning even if it represents less than 65 % of the population of the Union.

The double majority system will be used in the Council starting from the year 2014. However, a detailed analysis by Moberg (2010) shows that in this concrete case the 'double majority' system is not really double, as the per capita criterion (b') plays the dominant role here. In comparison with the Treaty of Nice, the voting power index will increase for the four largest states of the Union (Germany, France, the United Kingdom and Italy) and also for the smallest states. To understand this effect we shall analyze the voting system in which the voting weight of a given state is directly proportional to its population.

3.1 Voting Systems with Per Capita Criterion

The idea 'one citizen–one vote' looks so natural and appealing, that in several political debates one often did not care to analyze in detail its assumptions and all its consequences. It is somehow obvious that a minister representing a larger (if population is considered) state should have a larger weight during each voting in the EU Council. On the other hand, one needs to examine whether the voting weights of a minister in the Council should be proportional to the population he represents. It is clear that this would be very much the case, if one could assume that all citizens in each member state share the very same opinion in each case.

However, this assumption is obviously false, and nowadays we enjoy in Europe the freedom to express various opinions on every issue. Let us then formulate the question, how many citizens from his state each minister actually represents in an exemplary voting in the Council? Or to be more precise, how many voters from a given state with population N share in a certain case the opinion of their representative? We do not know!

Under the idealizing assumption that the minister always votes according to the will of the majority of citizens in his state, the answer can vary from $N/2$ to N. Therefore, the difference between the number of the citizens supporting the vote of their minister and the number of those who are against it can vary from 0 to N. In fact it will vary from case to case in this range, so an assumption that it is always proportional to N is false. This crucial issue, often overlooked in popular debates, causes problems with representativeness of a voting system based on the 'per capita' criterion.

There is no better way to tackle the problem as to rely on certain statistical assumptions and estimate the *average* number of 'satisfied citizens'. As such an analysis is performed later in this paper, we shall review here various arguments showing that a system with voting weights directly proportional to the population is advantageous to the largest states of the union.

Consider first a realistic example of a union of nine states: a large state A, with 80 millions of citizens and eight small states from B to I, with 10 millions each. Assume now that in a certain case the distribution of the opinion in the entire union is exactly polarized: in each state approximately 50 % of the population support the vote '*Yes*', while the other half is against. Assume now that the government of the large state is in position to establish exactly the will of the majority of citizens in their state (say it is the vote '*Yes*') and order its minister to vote accordingly. Thus the vote of this minister in the council will then be counted as a vote of 80 millions of citizens.

On the other hand, in the remaining states the probability that the majority of citizens support '*Yes*' is close to 50 %. Hence it is most likely that the votes of the ministers from the smaller states split as 4:4. Other outcomes: 5:3, 6:2, or 7:1 are less probable, but all of them result in the majority of the representative of the large state A. The outcome 8:0 is much less likely, so if we sum the votes of all nine ministers we see that the vote of the minister from the largest state will be decisive. Hence we have shown that the voting power of all citizens of the nine small states is negligible, and the decision for this model union is practically taken by the half of its population belonging to the largest state A. Even though in this example we concentrated on the 'per capita' criterion and did not take into account the other criterion, it is not difficult to come up with analogous examples which show that the largest states are privileged also in the double majority system. Similarly, the smallest states of the union benefit from the 'per state' criterion.

Let us have a look at the position of the minority in large states. In the above example the minority in the 80 million state can be as large as 40 million citizens, but their opinion will not influence the outcome of the voting, independently of the polarization of opinion in the remaining eight states. Thus one may conclude that in the voting system based on the 'per capita' criterion, the influence of the politicians representing the majority in a large state is enhanced at the expense of the minority in this state and the politicians representing the smaller states.

Last but not least, let us compare the maximal sizes of the minority, which can arise during any voting in an EU member state. In Luxembourg, with its population of about 400,000 people, the minority cannot exceed 200,000 citizens. On the other hand, in Germany, which is a much larger country, it is possible that the minority exceeds 41 millions of citizens, since the total population exceeds 82 millions. It is then fair to say, that, due to elections in smaller states, we know the opinion of citizens in these states with a better accuracy, than in larger members of the union. Thus, as in smaller states the number of misrepresented citizens is smaller, their votes in the EU Council should be weighted by larger weights than the vote of the largest states. This very idea is realized in the weighted voting system advocated by Penrose.

3.2 *Square Root Voting System of Penrose*

The Penrose system for the two-tier voting is based on the square root law reviewed in Sect. 2.1. Since the voting power of a citizen in state k with population N_k scales as $1/\sqrt{N_k}$, this factor will be compensated, if the voting power of each representative in the upper chamber will behave as $\sqrt{N_k}$. Only in this way the voting power of each citizen in every state of a union consisting of M states will be equal.

Although we know that the voting power of a minister in the Council needs not coincide with the weight of his vote, as a rough approximation let us put his weights w_k proportional to the square root of the population he represents, that is $w_k = \sqrt{N_k}/\sum_{i=1}^{M}\sqrt{N_i}$.

To see a possible impact of the change of the weights let us now return to the previous example of a union of one big state and eight small ones. As the state A is 8 times as large as each of the remaining states, its weight in the Penrose system will be $w_A = \sqrt{8}w_B$. As $\sqrt{8}$ exceeds 2 and is smaller then 3, we see that accepting the Penrose system will increase the role of the minority in the large state and the voting power of all smaller states. For instance, if the large state votes '*Yes*' and the votes in the eight states split as 2:6 or 1:7 in favor for '*No*', the decision will not be taken by the council, in contrast to the simple system with one 'per capita' criterion. There, we have assumed that the standard majority of weights is sufficient to form a winning coalition. If the threshold for the qualified majority is increased to 54 %, also the outcome 3:5 in favor for '*No*' in the smaller states suffices to block the decision taken in the large state.

This simple example shows that varying the quota for the qualified majority considerably influences the voting power, see also (Leech and Machover 2003; Machover 2010). The issue of the selection of the optimal quota will be analyzed in detail in the subsequent section. At this point, it is sufficient to add that in general it is possible to find such a level of the quota for which the voting power β_k of the kth state is proportional to $\sqrt{N_k}$ and, in consequence, the Penrose law is almost exactly fulfilled (Słomczyński and Życzkowski 2004, 2006).

Applying the square root voting system of Penrose combined with the optimal quota to the problem of the Council, one obtains a fair solution, in which every citizen in each member state of the Union has the same voting power, hence the same influence on the decisions taken by the Council. In this case, the voting power of each European state measured by the Banzhaf index scales as the square root of its population. This weighted voting system happens to give a larger voting power to the largest EU states (including Germany) than the Treaty of Nice but smaller than the double majority system. On the other hand, this system is more favorable to all middle size states then the double majority, so it is fair to consider it as a compromise solution. The square root voting system of Penrose is simple (one criterion only), transparent and efficient—the probability of forming a winning coalition is reasonably high. Furthermore, as discussed later, it can be easily adopted to any possible extension of the Union.

3.3 The Second Square Root Law of Morris

To provide an additional argument in favour of the square root weights of Felsenthal and Machover (1999), consider a model state of N citizens, of which a certain number k support a given legislation to be voted in the council. Assume that the representative of this state knows the opinion of his people and, according to the will of the majority, he votes '*Yes*' in the council if $k \geq N/2$. Then the number of citizens satisfied with his decision is k. The number $N - k$ of disappointed citizens compensates the same number of yes-votes, so the vote of the minister should effectively represent the *difference* between them, $w = k - (N - k) = 2k - N$. By our assumption concerning the majority this number is positive, but in general the effective weight of the vote of the representative should be $w = |2k - N|$.

Assume now that the votes of any of N citizens of the state are independent, and that both decisions are equally likely, so that $p = 1 - p = 1/2$. Thus, for the statistical analysis, we can use the Bernoulli scheme (3) and estimate the weight of the vote of the minister by the average using the Stirling approximation:

$$\langle w_N \rangle = \sum_{k=0}^{N} P_k |2k - N| = \sum_{k=0}^{N} \binom{N}{k} \frac{1}{2^N} |2k - N|$$
$$= \frac{\lfloor N/2 \rfloor + 1}{2^{N-1}} \binom{N}{\lfloor N/2 \rfloor + 1} \sim \sqrt{\frac{2N}{\pi}}. \tag{7}$$

Here $\lfloor x \rfloor$ denotes the largest integer not greater than x. This result provides another argument in favor of the weighted voting system of Penrose: Counting all citizens of a given state, we would attribute the weights of the representative proportionally to the population N he is supposed to represent. On the other hand, if we take into account the obvious fact that not all citizens in this state share the opinion of the government on a concrete issue and consider the average number of the majority of citizens which support his decision one should weight his vote proportionally to $\sqrt{N}$. From this fact one can deduce the *second square root law of Morriss* (Morriss 1987; Felsenthal and Machover 1998, 1999; Laruelle and Valenciano 2008) that states that the average number of misrepresented voters in the union is smallest if the weights are proportional to the square root of the population and quota is equal to 50 %, provided that the population of each member state is large enough. Simultaneously, in this situation, the total voting power of the union measured by the sum of the Banzhaf indices of all citizens in the union is maximal.

To illustrate the result consider a model union consisting of one large state with population of 49 millions, three medium states with 16 million each and three small with 1 million citizens. For simplicity assume that the double majority system and the Penrose system are based on the standard majority of 50 %. If the polarization of opinion in each state on a given issue is as in the table below, only 39 % of the population of the union is in favor of the legislative. However, under the rules of the double majority system the decision is taken (against the will of the vast majority!), what is not the case in the Penrose system, for which the coalition gains only 10 votes out of 22, so it fails to gather the required quota (see Table 1).

To qualitatively understand this result, consider the minister representing the largest country G with a population of 49 millions. In the double majority system

Table 1 *Case study* voting in the council of a model union of 7 members under a hypothetical distribution of population and voting preferences

State	*A*	*B*	*C*	*D*	*E*	*F*	*G*	Total	
Population (Million)	1	1	1	16	16	16	49	100	
Votes: Yes (Million)	2/3	2/3	2/3	4	4	4	25	39	
Votes: No (Million)	1/3	1/3	1/3	12	12	12	24	61	
State votes	1	1	1	0	0	0	1	4/7	Y
Minister's votes	1	1	1	0	0	0	49	52/100	Y
Square root weights	1	1	1	4	4	4	7	22	
Square root votes	1	1	1	0	0	0	7	10/22	N

Although 61 % of the total population of the union is against a legislative it will be taken by the council, if the rules of the double majority are used. The outcome of the voting according to the weighted voting system of Penrose correctly reflects the will of the majority in the union he uses his 49 votes against the will of 24 millions of inhabitants. By contrast, the minister of the small state A will misrepresent at most one half of the million of his compatriots. In other words, the precision in determining the will of all the citizens is largest in the smaller states, so the vote of their ministers should gain a higher weight than proportional to population, which is the case in the Penrose system.

4 Optimal Quota for Qualified Majority

Designing a voting system for the Council one needs to set the threshold for the qualified majority. In general, this quota can be treated as a free parameter of the system and is often considered as a number to be negotiated. For political reasons one usually requires that the voting system should be *moderately conservative*, so one considers the quota in the wide range from 55 to 75 %.

However, designing the voting system based on the theory of Penrose, one can find a way to obtain a single number as the optimal value of the quota. In order to assure that the voting powers of all citizens in the 'union' are equal one has to impose the requirement that the voting power of each member state should be proportional to the square root of the population of each state.

Let us analyze the problem of M members of the voting body, each representing a state with population N_i, $i = 1, \ldots, M$. Denote by w_i the voting weight attributed to each representative. We work with renormalized quantities, such that $\sum_{i=1}^{M} w_i = 1$. Assume that the decision of the voting body is taken, if the sum of the weights w_i of all members of the coalition exceeds the given quota q.

In the Penrose voting system one sets the voting weights proportional to the square root of the population of each state, $w_i \sim \sqrt{N_i}$ for $i = 1, , M$. For any level of the quota q one may compute numerically the power indices β_i. To characterize the overall representativeness of the voting system one may use various indices designed to quantify the resulting inequality in the distribution of power among citizens (Laruelle and Valenciano 2002). Analyzing the influence of the quota q for the average inequality of the voting power we are going to use the mean discrepancy Δ, defined as:

$$\Delta := \sqrt{\frac{1}{M}\sum_{i=1}^{M}(\beta_i - w_i)^2}, \tag{8}$$

If the discrepancy Δ is equal to zero, the voting power of each state is proportional to the square root of its population. Under the assumption that the Penrose law is fulfilled, in such a case the voting power of any citizen in each state is the same.

In practice, the coefficient Δ will not be exactly equal to zero, but one may try to minimize this quantity. The optimal quota q_* can be defined as the quota for which the discrepancy Δ is minimal. Let us note, however, that this definition works fine for the Banzhaf index, while the dependence of the Shapley–Shubik index (Shapley and Shubik 1954) on the quota does not exhibit such a minimum.

Table 2 Optimal quota q_n for the Council of the European Union of M member states compared with predictions q_{av} of the approximate formula (16) and the lower bound q_{min} given in (10)

M	25	27	28	29	...	$M \to \infty$
$q_n(\%)$	62.16	61.58	61.38	61.32	...	50.0
$q_{av}(\%)$	61.28	60.86	60.66	60.48	...	50.0
$q_{min}(\%)$	60.00	59.62	59.45	59.28	...	50.0

The calculations of the optimal quotas for the EU were based upon the Eurostat data on the distribution of population for the EU-25 (2004) and the EU-27 (2010). The extended variant EU-28 contains EU-27 and Croatia, while EU-29 includes also Iceland

Studying the problem for a concrete distribution of the population in the European Union, it was found (Słomczyński and Życzkowski 2004) that in these cases all M ratios β_i/w_i for $i = 1, \ldots, M$, plotted as a function of the quota q, cross approximately near a single point. In other words, the discrepancy Δ at this critical point q_* is negligible. Numerical analysis allows one to conclude that this optimal quota is approximately equal to 62 % for the EU-25 (Słomczyński and Życzkowski 2004). At this very level of the quota the voting system can be considered as optimal, since the voting power of all citizens becomes equal. Performing detailed calculations one needs to care to approximate the square root function with a sufficient accuracy, since the rounding effects may play a significant role (Kurth 2007).

It is worth to emphasize that in general the value of the optimal quota decreases with the number of member states. For instance, in the case of the EU-27 is is equal to 61.5 % (Życzkowski et al. 2006; Słomczyński and Życzkowski 2007), see Table 2. The optimal quota was also found for other voting bodies including various scenarios for an EU enlargement—see Leech and Aziz (2010). Note that the above results belong to the range of values of the quota for qualified majority, which are used in practice or recommended by experts.

4.1 Large Number of Member States and a Statistical Approximation

Further investigation has confirmed that the existence of such a critical point is not restricted to the concrete distribution of the population in European Union. On the contrary, it was reported for a model union containing M states with a random distribution of population (Słomczyński and Życzkowski 2004; Chang et al. 2006; Słomczyński and Życzkowski 2006). However, it seems unlikely that we can obtain an analytical expression for the optimal quota in such a general case. If the number of member states is large enough one may assume that the distribution of the sum of the weights is approximately Gaussian (Owen 1975; Feix et al. 2007; Słomczyński and Życzkowski 2007). Such an assumption allowed us to derive an explicit approximate formula for the optimal quota for the Penrose square root voting system (Słomczyński and Życzkowski 2007)

$$q_n \ := \ \frac{1}{2}\left(1+\frac{\sqrt{\sum_{i=1}^{M} N_i}}{\sum_{i=1}^{M}\sqrt{N_i}}\right), \tag{9}$$

where N_i denotes the population of the i–th state. In practice it occurs that already for $M = 25$ this approximation works fine and in the case of the EU-25 gives the optimal quota with an accuracy much better than one percent. Although the value of the optimal quota changes with M, the efficiency of the system, measured by the probability of forming the winning coalition, does not decrease if the union is enlarged. It was shown in Słomczyński and Życzkowski (2007) that, according to the central limit theorem, the efficiency of this system tends to approximately $15.9\,\%$ if $M \to \infty$.

It is not difficult to prove that for any fixed M the above expression attains its minimum if the population of each member state is the same, $N_i = const(i)$. In this way one obtains a lower bound for the optimal quota as a function of the number of states (Słomczyński and Życzkowski 2007):

$$q_{min} \ := \ \frac{1}{2}\left(1+\frac{1}{\sqrt{M}}\right). \tag{10}$$

Note that the above bound decreases with the number of the states forming the union as $1/\sqrt{M}$ to $50\,\%$. Such a behavior, reported in numerical analysis of the problem (Słomczyński and Życzkowski 2004; Chang et al. 2006; Słomczyński and Życzkowski 2006) is consistent with the so-called Penrose limit theorem—see Lindner and Machover (2004).

4.2 Optimal Quota Averaged over an Ensemble of Random States

Concrete values of the optimal quota obtained by finding numerically the minimum of the discrepancy (8) for the EU-25 and the EU-27 Słomczyński and Życzkowski (2004, 2006, 2010) are consistent, with an accuracy up to two per cent, with the data obtained numerically by averaging over a sample of random distribution of the populations of a fictitious union. This observation suggests that one can derive analytically an approximate formula for the optimal quota by averaging the explicit expression (9) over an ensemble of random populations N_i.

To perform such a task let us denote by x_i the relative population of a given state, $x_i = N_i / \sum_{i=1}^{M} N_i$. Since $\sqrt{N_i}/\sqrt{\sum_{i=1}^{M} N_i} = \sqrt{x_i}$ one can rewrite expression (9) in the new variables to obtain

$$q_n(\overrightarrow{x}) = \frac{1}{2}\left(1 + \frac{1}{\sum_{i=1}^{M}\sqrt{N_i}/\sqrt{\sum_{i=1}^{M}N_i}}\right) = \frac{1}{2}\left(1 + \frac{1}{\sum_{i=1}^{M}\sqrt{x_i}}\right). \tag{11}$$

By construction, $\overrightarrow{x} = (x_1, \ldots, x_M)$ forms a probability vector with $x_i \geq 0$ and $\sum_{i=1}^{M} x_i = 1$. Hence the entire distribution of the population of the union is characterized by the M-point probability vector $\overrightarrow{x}$, which lives in an $(M-1)$ dimensional simplex Δ_M. Without any additional knowledge about this vector we can assume that it is distributed uniformly on the simplex,

$$P_D(x_1, \ldots, x_M) = \frac{1}{(M-1)!}\,\delta\left(1 - \sum_{i=1}^{M} x_i\right). \tag{12}$$

Technically it is a particular case of the *Dirichlet distribution*, written $P_D(\overrightarrow{x})$, with the Dirichlet parameter set to unity.

In order to get a concrete result one should then average expression (11) with the flat probability distribution (12). Result of such a calculation can be roughly approximated by substituting M-fold mean value over the Dirichlet measure, $M\langle\sqrt{x}\rangle_D$, instead of the sum into the denominator of the correction term in (11),

$$q_{av}(M) := \langle q_n\rangle_D \approx \frac{1}{2}\left(1 + \frac{1}{M\langle\sqrt{x}\rangle_D}\right). \tag{13}$$

The mean square root of a component of the vector $\overrightarrow{x}$ is given by an integral with respect to the Dirichlet distribution

$$\langle\sqrt{x}\rangle_D = \int_{\Delta_M} \sqrt{x_1}\, P_D(x_1, \ldots, x_M)\, dx_1 \cdots dx_M. \tag{14}$$

Instead of evaluating this integral directly, we shall rely on some simple fact from the physical literature. It is well known that the distribution of the squared absolute values of an expansion of a random state in an M-dimensional complex Hilbert space is given just by the flat Dirichlet distribution (see e.g. Bengtsson and Życzkowski 2006). In general, all moments of such a distribution where computed by Jones (1991). The average square root is obtained by taking his expression from Jones and setting $d = M$, $l = 1$, $v = 2$ and $\beta = 1/2$. This gives the required average

$$\langle\sqrt{x}\rangle_D = \frac{\Gamma(M)\,\Gamma(3/2)}{\Gamma(M+1/2)} \sim \frac{\sqrt{\pi}}{2\sqrt{M}}. \tag{15}$$

Here Γ denotes the Euler gamma function and the last step follows from its Stirling approximation. Substituting the average $\langle\sqrt{x}\rangle_D$ into (13) we arrive at a compact expression

$$q_{av}(M) \approx \frac{1}{2} + \frac{1}{\sqrt{\pi M}} = \frac{1}{2}\left(1 + \frac{2}{\sqrt{\pi}}\frac{1}{\sqrt{M}}\right). \tag{16}$$

This approximate formula for the mean optimal quota for the Penrose voting system in a union of M random states constitutes the central result of this work. Note that this expression is averaged over all possible distributions of populations in the union, so it depends only on the size M of the union and on the form of averaging. The formula has a similar structure as the lower bound (10), but the correction term is enhanced by the factor $2/\sqrt{\pi} \approx 1.128$. In some analogy to the famous *Buffon's needle (or noodle) problem* (Ramaley 1969), the final result contains the number π—it appears in (16) as a consequence of using the normal approximation. The key advantage of the result (16) is due to its simplicity. Therefore, it can be useful in a practical case, if the size M of the voting body is fixed, but the weights of the voters (e.g. the populations in the EU) vary.

5 Concluding Remarks

In this work we review various arguments leading to the weighted voting system based upon the square root law of Penrose. However, the key result consists in an approximate formula for the mean optimal threshold of the qualified majority. It depends only on the number M of the states in the union, since the actual distribution of the population is averaged out.

Making use of this result we are in a position to propose a simplified voting system. The system consists of a single criterion only and is determined by the following two rules:

(1) Each member of the voting body of size M is attributed his voting weight proportional to the square root of the population he represents.
(2) The decision of the voting body is taken if the sum of the weights of members of a coalition exceeds the critical quota $q = 1/2 + 1/\sqrt{\pi M}$.

This voting system is based on a single criterion. Furthermore, the quota depends on the number of players only, but not on the particular distribution of weights of the individual players. This feature can be considered as an advantage in a realistic case, if the distribution of the population changes in time. The system proposed is objective and it cannot a priori handicap a given member of the voting body. The quota for qualified majority is considerably larger than 50 % for any size of the voting body of a practical interest. Thus the voting system is moderately conservative, as it should be. If the distribution of the population is known and one may assume that it is invariant in time, one may use a modified rule (2**'**) and set the optimal quota according to the more precise formula (9).

Furthermore, the system is transparent: the voting power of each member of the voting body is up to a high accuracy proportional to his voting weight. However, as a crucial advantage of the proposed voting system we would like to emphasize its extendibility: if the size M of the voting body changes, all one needs to do is to set the voting weights according to the square root law and adjust the quota q according to the rule (2'). Moreover, for a fixed number of players, the system does not depend on the particular distribution of weights. This feature is specially relevant for voting bodies in corporate management for which the voting weights may vary frequently.

It is our pleasure to thank W. Kirsch, M. Machover, F. Pukelsheim for fruitful discussions and to E. Ratzer for helpful correspondence. We are obliged to the anonymous referee for numerous comments and suggestions which allowed us to improve the article. Financial support by the European grant COCOS is gratefully acknowledged.

References

Ade, F. (2006). Decision making in Europe: Were Spain and Poland right to stop the constitution in December 2003? (Preprint 2006), from http://congress.utu.fi/epcs2006/docs/D3_ade.pdf

Algaba, E., Bilbao, J. R., & Fernandez, J. R. (2007). The distribution of power in the European constitution. *European Journal of Operational Research*, *176*, 1752–1766.

Andjiga, N.-G., Chantreuil, F., & Leppeley, D. (2003). La mesure du pouvoir de vote. *Mathematical Social Sciences*, *163*, 111–145.

Baldwin, R. E., & Widgrén, M. (2004). Winners and losers under various dual majority rules for the EU Council of Ministers. CEPR Discussion Paper No. 4450, Centre for European Policy Studies, Brussels 2004, from http://www.cepr.org/pubs/dps/DP4450.asp

Banzhaf, J. F. (1965). Weighted voting does not work: A mathematical analysis. *Rutgers Law Review*, *19*, 317–343.

Bârsan-Pipu, N., & Tache, I. (2009). An analysis of EU voting procedures in the enlargement context. *International Advances in Economic Research*, *15*, 393–408.

Beisbart, C., Bovens, L., & Hartmann, S. (2005). A utilitarian assessment of alternative decision rules in the Council of Ministers. *European Union Politics*, *6*, 395–419.

Bengtsson, I., & Życzkowski, K. (2006). *Geometry of quantum states*. Cambridge: Cambridge UP.

Chang, P.-L., Chua, V. C. H., & Machover, M. (2006). LS Penrose's limit theorem: tests by simulation. *Mathematical Social Sciences*, *51*, 90–106.

Feix, M. R., Lepelley, D., Merlin, V., & Rouet, J. L. (2007). On the voting power of an alliance and the subsequent power of its members. *Social Choice and Welfare*, *28*, 181–207.

Felsenthal, D. S., & Machover, M. (1997). The weighted voting rule in the EUs Council of Ministers, 1958–95: Intentions and outcomes. *Electoral Studies*, *16*, 33–47.

Felsenthal, D. S., & Machover, M. (1998). *Measurement of voting power: Theory and practice, problems and paradoxes*. Cheltenham: Edward Elgar.

Felsenthal, D. S., & Machover, M. (1999). Minimizing the mean majority deficit: The second square-root rule. *Mathematical Social Science*, *37*, 25–37.

Felsenthal, D. S., & Machover, M. (2001). Treaty of nice and qualified majority voting. *Social Choice and Welfare*, *18*, 431–464.

Gelman, A., Katz, J. M., & Tuerlinckx, F. (2002). The mathematics and statitics of voting power. *Statistical Science*, *17*, 420–435.

Gelman, A., Katz, J. M., & Bafumi, J. (2004). Standard voting power indexes do not work: an empirical analysis. *British Journal of Political Science*, *34*, 657–674.

Hosli, M. O. (2008). Council decision rules and European Union constitutional design. *AUCO Czech Economic Review*, *2*, 76–96.

Jones, K. R. W. (1991). Riemann-Liouville fractional integration and reduced distributions on hyperspheres. *Journal of Physics A*, *24*, 1237–1244.

Kirsch, W., Słomczyński, W., & Życzkowski, K. (2007). Getting the votes right. *European Voice*, *3–9*, 12.

Kirsch, W. (2007). On Penrose's square-root law and beyond. *Homo Oeconomicus*, *24*, 357–380.

Kirsch, W. (2010). The distribution of power in the Council of Ministers of the European Union. In M. Cichocki & K. Życzkowski (Eds.), *Institutional design and voting power in the European Union* (pp. 93–107). Farnham: Ashgate Publishing Group.

Kurth, M. (2007). Square root voting in the Council of the European Union: Rounding effects and the jagiellonian compromise, (Preprint math.GM 0712.2699).

Laruelle, A., & Widgrén, M. (1998). Is the allocation of voting power among the EU states fair? *Public Choice*, *94*, 317–339.

Laruelle, A., & Valenciano, F. (2002). Inequality among EU citizens in the EU's Council decision procedure. *European Journal of Political Economy*, *18*, 475–498.

Laruelle, A., & Valenciano, F. (2008). *Voting and collective decision-making. Bargaining and power*. Cambridge: Cambridge UP.

Leech, D. (2002). Designing the voting system for the Council of the EU. *Public Choice*, *113*, 437–464.

Leech, D., & Machover, M. (2003). Qualified majority voting: The effect of the quota. In M. Holler, et al. (Eds.), *European Governance, Jahrbuch für Neue Politische Ökonomie* (pp. 127–143). Mohr Siebeck: Tübingen.

Leech, D., & Aziz, H. (2010). The double majority voting rule of the EU reform treaty as a democratic ideal for an enlarging union: An appraisal using voting power analysis. In M. Cichocki & K. Życzkowski (Eds.), *Institutional design and voting power in the European Union* (pp. 59–73). Farnham: Ashgate Publishing Group.

Lindner, I., & Machover, M. (2004). LS Penrose's limit theorem: proof of some special cases. *Mathematical Social Sciences*, *47*, 37–49.

Machover, M. (2010). Penrose's square root rule and the EU Council of the Ministers: Significance of the quota. In M. Cichocki & K. Życzkowski (Eds.), *Institutional design and voting power in the European Union* (pp. 35–42). Farnham: Ashgate Publishing Group.

Moberg, A. (2010). Is the double majority really double? The voting rules in the Lisbon Treaty. In M. Cichocki & K. Życzkowski (Eds.), *Institutional design and voting power in the European Union* (pp. 19–34). Farnham: Ashgate Publishing Group.

Morriss, P. (1987) *Power: A philosophical analysis* (2nd ed. 2002). Manchester UP: Manchester.

Owen, G. (1975). Multilinear extensions and the Banzhaf value. *Naval Research Logistics Quaterly*, *22*, 741–750.

Pajala, A., & Widgrén, M. (2004). A priori versus empirical voting power in the EU Council of Ministers. *European Union Politics*, *5*, 73–97.

Penrose, L. S. (1946). The elementary statistics of majority voting. *Journal of the Royal Statistical Society*, *109*, 53–57.

Penrose, L. S. (1952). *On the objective study of crowd behaviour*. London: H.K. Lewis & Co.

Pöppe, Ch. Die Quadratwurzel, das Irrationale und der Tod, *Spektrum der Wissenschaft*, August 2007, pp. 102–105.

Pukelsheim, F. (2007). Der Jagiellonische Kompromiss. *Neue Züricher Zeitung*, 20 Juni 2007.

Pukelsheim, F. (2010). Putting citizens first: Representation and power in the European Union in institutional design and voting power in the European Union. In M. Cichocki, & K. Życzkowski (Eds.), (pp. 235–253). Farnham: Ashgate Publishing Group

Ramaley, J. F. (1969). Buffon's noodle problem. *American Mathematical Monthly*, *76*, 916–918.

Shapley, L. S., & Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American Science Review*, *48*, 787–792.

Słomczyński, W., & Życzkowski, K. (2004) Voting in the European Union: The square root system of penrose and a critical point. preprint cond-mat.0405396.

Słomczyński, W., & Życzkowski, K. (2006). Penrose voting system and optimal quota. *Acta Physica Polonica B*, *37*, 3133–3143.

Słomczyński, W., & Życzkowski, K. (2007). From a toy model to the double square root voting system. *Homo Oeconomicus*, *24*, 381–399.

Słomczyński, W., & Życzkowski, K. (2010). Jagiellonian compromise—an alternative voting system for the Council of the European Union. In M. Cichocki & K. Życzkowski (Eds.), *Institutional design and voting power in the European Union* (pp. 43–57). Farnham: Ashgate Publishing Group.

Życzkowski, K., Słomczyński, W., & Zastawniak, T. (2006). Physics for fairer voting. *Physics World*, *19*, 35–37.

The QM Rule in the Nice and Lisbon Treaties: Future Projections

Dan S. Felsenthal and Moshé Machover

1 Introductory Remarks

The Treaty of Lisbon, amending the Treaty on European Union and the Treaty Establishing the European Community, was signed in Lisbon by all 27 EU members on 13 December 2007. The aim of the treaty, as stated in its preamble, is '... to complete the process started by the Treaty of Amsterdam and by the Treaty of Nice with a view to enhancing the efficiency and democratic legitimacy of the Union and to improving the coherence of its action';.[1] Purportedly towards this aim, the treaty incorporates a new qualified majority (QM) decision rule for the EU Council of Ministers (CM).[2] Since the treaty has been ratified by all EU members, this rule, as cited below, is due to take effect not earlier than 1 November 2014 and not later than 31 March 2017. Until then, the QM rule contained in the Nice Treaty (Treaty of Nice 2001) will remain in force. The new rule states:

An earlier version of this chapter was published in *Homo Oeconomicus* **26**(3/4), 2009.

[1] See (Treaty of Lisbon (2007), p. 3).
[2] This rule was first adopted at the Brussels IGC, 17–18 June 2004 (See CIG 85/04 2004), and was included in the proposed EU Constitution, which failed to be ratified and was abandoned. Subsequently, the same rule was confirmed on 23 June 2007 by the Council of the European Union (the 'EU Summit'), also held in Brussels. (See 11177/1/07 REV 1 (2007)).

D. S. Felsenthal
University of Haifa, Haifa, Israel
e-mail: dfelsenthal@poli.haifa.ac.il

D. S. Felsenthal · M. Machover (✉)
Centre for the Philosophy of the Natural and Social Sciences,
London School of Economics and Political Science, Lakatos Building,
Houghton Street, London WC2A 2AE, UK
e-mail: moshe.machover@kcl.ac.uk

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_31,

The Treaty on European Union shall be amended [such that]...

17) An Article 9c shall be inserted:

Article 9c

1. The Council shall, jointly with the European Parliament, exercise legislative and budgetary functions. It shall carry out policy-making and coordinating functions as laid down in the Treaties.
2. The Council shall consist of a representative of each Member State at ministerial level, who may commit the government of the Member State in question and cast its vote.
3. The Council shall act by qualified majority except where the Treaties provide otherwise.
4. As from 1 November 2014, a qualified majority shall be defined as at least 55 % of the members of the Council, comprising at least 15 of them and representing Member States comprising at least 65 % of the population of the Union. A blocking minority must include at least four Council members, failing which the qualified majority shall be deemed attained. The other arrangements governing the qualified majority are laid down in Article 205(2) of the Treaty on the Functioning of the European Union.
5. The transitional provisions relating to the definition of the qualified majority which shall be applicable until 31 October 2014 and those which shall be applicable from 1 November 2014 to 31 March 2017 are laid down in the Protocol on transitional provision.[3]

This decision rule depends *explicitly* on the size of population of the member-states. Thus, the number and composition of coalitions able to pass or block a decision of the CM, as well as the voting powers of the member-states (and other related quantities) will be automatically affected by demographic changes. Strictly speaking, it is not a single fixed rule, but a variable rule that depends not only on the number of member-states but also on their changing populations. Formally, the same holds also for the Nice QM rule, which is currently in force. However, as we showed in Felsenthal and Machover (2001), the effect of the population clause in the latter rule is rather insignificant if not negligible.[4] So the new QM rule is the first in the history of the EU whose functioning can be affected *significantly* by changes in population size.[5]

In the present chapter we describe and analyse the effects on the distribution of voting powers and related quantities that would result from the demographic changes forecast by Eurostat (2008) for the period stretching from the year 2008 to the year 2060.

---

3 See (Treaty of Lisbon (2007), p. 18). Article 9c contains four additional clauses which we have not cited here as they are not directly pertinent to this article.

4 The main—and dominant—clause in the Nice QM rule assigned voting weights to the member-states; these weights took account of population sizes as they were at the time (2000), but were to remain fixed henceforth.

5 In view of this fact it is rather strange that—as far as we know—the EU does not have a uniform definition of the 'population' of each member-state, and a legally binding procedure ascertaining its size at synchronized regular intervals.

We confine ourselves exclusively to the distribution of voting power within the CM. We express no view as to the importance of this issue in the general governance of the EU or to the future of the EU. For differing views on these questions, see e.g. Nurmi (2008) and Widgrén (2008).

We did the computations of voting powers for the years 2008–2014 at three-year intervals, and for the years 2015–2060 at five-year intervals, using (Eurostat 2008) population forecasts. Our calculations for 2008, 2011 and 2014 are done under the Nice QM rule, whereas from 2015 we assume the new Lisbon Treaty rule.

For simplicity we assume that the current EU membership of 27 states will be unchanged throughout this period.

We find that the clause excluding blocking coalitions with less than four members rules out during the period 2015–2060 only 10–12 coalitions of three member-states, whose populations comprise more than 35 % of the total, and therefore would other-wise be able to block. In 2015 there are 10 such coalitions, namely:

{Germany, UK, France},
{Germany, UK, Italy},
{Germany, UK, Spain},
{Germany, UK, Poland},
{Germany, France, Italy},
{Germany, France, Spain},
{Germany, France, Poland},
{Germany, Italy, Spain},
{Germany, Italy, Poland},
{UK, France, Italy}.

From 2020 onwards there is an additional (11th) such coalition: {UK, France, Spain}. As of 2025 the coalition {Germany, Italy, Poland} is expected to consist of fewer than 35 % of the EU population, so between 2025 and 2040 there are expected to be again only 10 three-member coalitions comprising more than 35 % of the EU population. As of 2045 the coalition {Germany, France, Poland} is expected to consist of fewer than 35 % of the EU population, but because the coalitions {UK, Italy, Spain} and {France, Italy, Spain} are expected to comprise more than 35 % of the EU population, the total number of three-member coalitions comprising more than 35 % of the EU population is expected to be again 11 during the period 2045–2055. This total number is expected to grow to 12 as of 2055 when the coalition {UK, France, Poland} is expected to comprise more than 35 % of the EU population.

We have taken these exceptional coalitions into account in our calculations; but in any case their effect on voting powers and related quantities is negligible.

The results of our calculations are presented in Tables 1–7. The general structure of these tables is described in Sect. 3. The meaning of the various measures and parameters presented in the tables is outlined in Sect. 2. Our conclusions are presented in Sect. 4.

2 Explanations

In this section we explain the meaning of the measures used in the following two sections and the criteria used in our assessment of QM decision rules. Our method here is largely the same as in Felsenthal and Machover (2004a), where the reader can find some further explanatory details.

2.1 Voting Power: Absolute, Relative and Negative

Each of the three series of values ψ, β and γ conveys information on a different aspect of voting power.

Penrose's measure ψ is an objective measure of *absolute* a priori voting power; its value for a given voter quantifies the amount of influence over the outcomes of divisions that the voter derives from the decision rule itself.

Thus, if the value of ψ for a member-state is higher under decision rule $\mathcal{U}$ than under $\mathcal{V}$, it follows that the position of that member-state is objectively better—in the sense of having more influence—under $\mathcal{U}$ than under $\mathcal{V}$. The importance of ψ for comparing the position of a given voter under different decision rules is not sufficiently appreciated even by some academic commentators.

Politicians are obviously interested in comparing the *relative* position of their country with those of other member-states, especially ones whose populations are close in size to their own. As far as we know, they do not employ the precise scientific measure of a priori relative voting power, the Banzhaf index β, which is obtained from ψ by normalization. Instead, they look at the voting weights, which can give a rough—and often quite imprecise—idea about relative voting power.

Another aspect of voting power in which politicians are keenly interested is negative or blocking power—the ability to help block an act that they oppose. Of course, this does not mean that they have more than a vague notion as to how to quantify this power.

Absolute voting power, as measured by ψ, is the voter's ability to help secure a favourable outcome in a division. This can be resolved into two component parts: the power to help secure a positive outcome, approval of an act that the voter supports; and the power to help secure a negative outcome, blocking of an act that the voter opposes. These two components are quantified by the Coleman measures γ^* and γ, respectively. From a purely objective, disinterested viewpoint, both are equally important; and indeed ψ is a symmetric combination of γ^* and γ.[6]

[6] In fact, ψ is their *harmonic mean*. For further details see (Felsenthal and Machover (1998), pp. 49–51).

However, for rather obvious political reasons, EU practitioners are much more concerned about negative voting power than about its positive counterpart.[7]

So in this article we present all three sets of data about the QM rules under consideration: ψ as an objective measure of absolute voting power; as well as β and γ, which quantify aspects of voting power that are of particular concern to practitioners.

2.2 Democratic Legitimacy

The CM can be regarded as the upper tier of a two-tier decision-making structure: if we assume that each minister votes in the CM according to the majority opinion in his or her country, then the citizens of the EU are seen as indirect voters, voting via their respective representatives at the CM. The criteria considered under the present heading are *equitability* and *adherence to majority rule*. These address different aspects of the functioning of the CM as the upper tier of the two-tier structure.

As explained elsewhere (see (Felsenthal and Machover (1998), pp. 66–67)), a perfectly equitable decision rule for the CM—in the sense of equalizing the indirect a priori voting powers of all EU citizens across all member-states—would give each member-state voting power proportional to the square root of its population size. (This is Penrose's Square-Root Rule.) So under such a decision rule the value β_i of β for member-state i would equal

$$\hat{\beta}_i := \frac{\sqrt{p_i}}{\sum_{j=1}^{27} \sqrt{p_j}},$$

where p_i is the population of member-state i. The Quotient is defined as the actual value of β divided by the 'equitable ideal' $\hat{\beta}$. In other words, the value Q_i of the Quotient for member-state i is

$$Q_i := \frac{\beta_i \sum_{j=1}^{27} \sqrt{p_j}}{\sqrt{p_i}}.$$

The amount by which the Quotient for a given member-state exceeds or falls short of 1 indicates the amount by which the voting power of this member-state exceeds or falls short of what it should have got under an equitable distribution of the same amount of total voting power.

---

[7] For reasons of internal national politics, a government normally considers it more important to be able to block a CM act that it opposes than to secure approval of an act it favours. Also, a government that finds itself in a position where it would be able to block a CM act may use this as a bargaining chip: agree to vote for the act in exchange for concession on matters that may or may not be related to that act.

In order to assess the degree to which a given rule is equitable, we therefore gauge how close its 27 β values are to the ideal presented by the corresponding $\hat{\beta}$ values. For this purpose we use three synoptic parameters. All three are given in percentage terms—hence the coefficient 100 in their definitions:

D This is the widely used *index of distortion*. It is defined as:

$$D := 100 \sum_{i=1}^{27} \frac{|\beta_i - \hat{\beta}_i|}{2}.$$

The smallest possible value of D is 0 and its greatest possible value is 100. The *smaller* the value of D, the closer the overall fit between the β_i and $\hat{\beta}_i$.

max|d| *Maximal relative deviation*. It is defined as:

$$\max |d| := 100 \max_i |Q_i - 1|.$$

ran(d) *Range of relative deviations*. It is defined as:

$$ran(d) := 100(\max_i Q_i - \min_i Q_i).$$

D is a measure of the *overall* discrepancy between the 27 β values and the corresponding $\hat{\beta}$ values. Thus it can serve as a measure of the overall equitability of the decision rule in question. On the other hand max $|d|$ and $ran(d)$ quantify the most extreme *individual* deviations of the given rule from equitability.

We now turn to our criterion of adherence to majority rule. In any non-trivial two-tier decision-making structure it can happen that the decision at the upper tier (in our case: the CM) goes against the majority view of the lower-tier indirect voters (in our case: the citizens of the EU at large). In a case where this happens—that is, the CM approves an act that is opposed by a majority of EU citizens, or blocks an act that is supported by a majority of the citizens—the margin by which the majority that opposes the decision exceeds the minority that supports it is the *majority deficit* of this decision. In a case where the majority of citizens support the CM decision the majority deficit is 0. The majority deficit can be regarded as a random variable (taking only non-negative integer values), whose distribution depends on the decision rule of the CM. The mean value (mathematical expectation) of this random variable is the *mean majority deficit* (MMD).[8] The larger the MMD, the further the CM decision rule is from the majoritarian ideal.

[8] For details, see (Felsenthal and Machover (1998), pp. 60–61).

Table 1 ψ values by country and year

Country	2008	2011	2014	2015	2020	2025
Austria	0.012988	0.012988	0.012988	0.043489	0.043485	0.043985
Belgium	0.015474	0.015474	0.015474	0.048533	0.048649	0.049091
Bulgaria	0.012988	0.012988	0.012988	0.041041	0.040462	0.040225
Cyprus	0.005251	0.005251	0.005251	0.027819	0.028011	0.028312
Czech Rep.	0.015474	0.015474	0.015474	0.047321	0.047053	0.047121
Denmark	0.009160	0.009160	0.009160	0.037401	0.037463	0.037654
Estonia	0.005251	0.005251	0.005251	0.028635	0.028616	0.028912
Finland	0.009160	0.009160	0.009160	0.036994	0.036866	0.037258
France	0.032688	0.032688	0.032688	0.153603	0.154926	0.156792
Germany	0.032688	0.032688	0.032688	0.193354	0.190137	0.187819
Greece	0.015474	0.015474	0.015474	0.049341	0.049244	0.049287
Hungary	0.015474	0.015474	0.015474	0.046310	0.045857	0.045743
Ireland	0.009160	0.009160	0.009160	0.036386	0.036866	0.037654
Italy	0.032688	0.032688	0.032688	0.146465	0.146003	0.146144
Latvia	0.005251	0.005251	0.005251	0.030471	0.030426	0.030509
Lithuania	0.009160	0.009160	0.009160	0.032734	0.032447	0.032662
Luxembourg	0.005251	0.005251	0.005251	0.026998	0.027202	0.027508
Malta	0.003957	0.003957	0.003957	0.026792	0.026797	0.027115
Netherlands	0.016691	0.016691	0.016706	0.059931	0.059884	0.060161
Poland	0.031163	0.031163	0.031163	0.091485	0.088740	0.086665
Portugal	0.015474	0.015474	0.015474	0.048128	0.048248	0.048500
Romania	0.017888	0.017888	0.017888	0.069123	0.067860	0.066979
Slovakia	0.009160	0.009160	0.009160	0.036994	0.037065	0.037062
Slovenia	0.005251	0.005251	0.005251	0.030268	0.030226	0.030308
Spain	0.031164	0.031164	0.031164	0.121403	0.123819	0.125610
Sweden	0.012988	0.012988	0.012988	0.045502	0.045857	0.046337
UK	0.032688	0.032688	0.032688	0.152719	0.155145	0.158323

Country	2030	2035	2040	2045	2050	2055	2060
Austria	0.043682	0.043757	0.043676	0.043584	0.043632	0.043653	0.043701
Belgium	0.048693	0.049053	0.049057	0.049039	0.049189	0.049143	0.049507
Bulgaria	0.039432	0.038828	0.038460	0.038082	0.037835	0.037571	0.037466
Cyprus	0.028344	0.028512	0.028884	0.029030	0.029438	0.029618	0.029910
Czech Rep.	0.046379	0.046025	0.045718	0.045222	0.045072	0.044721	0.044589
Denmark	0.037493	0.037686	0.037709	0.037714	0.037835	0.037926	0.038182
Estonia	0.028736	0.028705	0.028884	0.029030	0.029254	0.029437	0.029551
Finland	0.037106	0.037113	0.036964	0.036978	0.036924	0.037029	0.037286
France	0.157868	0.159586	0.161329	0.162820	0.164277	0.165857	0.167787
Germany	0.184027	0.180880	0.177720	0.174274	0.171378	0.168279	0.165769
Greece	0.048884	0.048486	0.048315	0.047951	0.047760	0.047555	0.047406
Hungary	0.045030	0.044514	0.044233	0.043764	0.043632	0.043294	0.043169
Ireland	0.037689	0.038065	0.038275	0.038635	0.038926	0.039363	0.039791

(continued)

Table 1 (continued)

Country	2030	2035	2040	2045	2050	2055	2060
Italy	0.145472	0.145682	0.145933	0.145970	0.145742	0.145258	0.144639
Latvia	0.030102	0.030240	0.030206	0.030328	0.030355	0.030344	0.030631
Lithuania	0.032218	0.032146	0.032089	0.031998	0.032005	0.031982	0.032078
Luxembourg	0.027365	0.027549	0.027939	0.028102	0.028338	0.028531	0.028826
Malta	0.026972	0.027163	0.027374	0.027544	0.027785	0.027984	0.028284
Netherlands	0.059489	0.059205	0.058607	0.058026	0.057493	0.056963	0.056915
Poland	0.083438	0.080810	0.078312	0.076061	0.074410	0.073120	0.072544
Portugal	0.048114	0.048106	0.048131	0.047951	0.047760	0.047731	0.047758
Romania	0.064921	0.063545	0.062273	0.060878	0.059575	0.058503	0.057423
Slovakia	0.036529	0.036355	0.036216	0.036052	0.036014	0.035763	0.035672
Slovenia	0.030102	0.030240	0.030393	0.030328	0.030535	0.030523	0.030812
Spain	0.126278	0.127218	0.128344	0.129239	0.129945	0.130198	0.129960
Sweden	0.046188	0.046215	0.046277	0.046314	0.046506	0.046673	0.047056
UK	0.160443	0.163180	0.165700	0.168355	0.171378	0.174274	0.177542

2.3 Efficiency

The criteria we consider under this heading address the functioning of the CM as a decision-making body in its own right rather than as part of a two-tier structure.

The [absolute] sensitivity of a decision rule is the sum of the voting powers (as measured by ψ) of all members of the CM. It measures the degree to which the CM collectively is empowered as a decision-making body, the ease with which an average member can make a difference to the outcome of a division. It is thus a good indicator of efficiency.

The *relative sensitivity index*, denoted by S, measures the sensitivity of the given rule on a logarithmic scale, on which $S = 0$ holds for the least sensitive rule (unanimity) with the same number of voters, and $S = 1$ holds for the most sensitive rule (the ordinary majority rule) with that number of voters.[9]

The second criterion under the present heading is that of *compliance*. A direct measure of this is Coleman's 'power of the collectivity to act', which is simply the a priori probability A of an act being approved rather than blocked.

A measures the compliance of a decision rule, the ease with which a positive outcome is approved. But it is often instructive to look at its reverse, so to speak: the resistance of a decision rule to approving an act. A convenient measure of this is the *resistance coefficient* R.[10] For proper decision rules, the least value of R is 0 (attained for a simple majority rule with an odd number of voters) and its maximal value is 1 (attained by the unanimity rule).

[9] For further details see (Felsenthal and Machover (1998), p. 61).

[10] For further details see (Felsenthal and Machover (1998), p. 62).

Table 2 100β values by country and year

Country	2008	2011	2014	2015	2020	2025
Austria	3.0924	3.0924	3.0923	2.5444	2.5469	2.5666
Belgium	3.6843	3.6843	3.6842	2.8394	2.8494	2.8645
Bulgaria	3.0924	3.0924	3.0923	2.4011	2.3699	2.3472
Cyprus	1.2502	1.2502	1.2501	1.6276	1.6406	1.6521
Czech Rep.	3.6843	3.6843	3.6842	2.7686	2.7559	2.7496
Denmark	2.1809	2.1808	2.1808	2.1882	2.1942	2.1972
Estonia	1.2502	1.2502	1.2501	1.6753	1.6760	1.6871
Finland	2.1809	2.1808	2.1808	2.1643	2.1592	2.1741
France	7.7828	7.7828	7.7825	8.9866	9.0740	9.1491
Germany	7.7828	7.7828	7.7826	11.3131	11.1363	10.9596
Greece	3.6843	3.6843	3.6842	2.8867	2.8842	2.8760
Hungary	3.6843	3.6843	3.6842	2.7094	2.6859	2.6692
Ireland	2.1809	2.1808	2.1808	2.1288	2.1592	2.1972
Italy	7.7827	7.7827	7.7825	8.5690	8.5514	8.5278
Latvia	1.2502	1.2502	1.2501	1.7827	1.7821	1.7802
Lithuania	2.1809	2.1808	2.1808	1.9151	1.9005	1.9059
Luxembourg	1.2502	1.2502	1.2501	1.5795	1.5932	1.6051
Malta	0.9422	0.9422	0.9422	1.5675	1.5695	1.5822
Netherlands	3.9740	3.9740	3.9775	3.5063	3.5074	3.5105
Poland	7.4198	7.4198	7.4195	5.3524	5.1975	5.0571
Portugal	3.6843	3.6843	3.6842	2.8158	2.8259	2.8301
Romania	4.2591	4.2591	4.2589	4.0441	3.9746	3.9084
Slovakia	2.1809	2.1808	2.1808	2.1643	2.1709	2.1626
Slovenia	1.2502	1.2502	1.2501	1.7709	1.7703	1.7685
Spain	7.4199	7.4199	7.4197	7.1028	7.2521	7.3296
Sweden	3.0924	3.0924	3.0923	2.6621	2.6859	2.7038
UK	7.7828	7.7828	7.7825	8.9349	9.0869	9.2385
Total	100.0003	99.9998	100.0003	100.0009	99.9999	99.9998

Country	2030	2035	2040	2045	2050	2055	2060
Austria	2.5680	2.5756	2.5737	2.5740	2.5772	2.5811	2.5794
Belgium	2.8626	2.8874	2.8908	2.8961	2.9054	2.9057	2.9221
Bulgaria	2.3182	2.2855	2.2663	2.249	2.2348	2.2214	2.2113
Cyprus	1.6663	1.6783	1.7021	1.7144	1.7388	1.7512	1.7654
Czech Rep.	2.7266	2.7092	2.6940	2.6707	2.6623	2.6442	2.6318
Denmark	2.2042	2.2183	2.2221	2.2273	2.2348	2.2424	2.2536
Estonia	1.6894	1.6897	1.7021	1.7144	1.7279	1.7405	1.7442
Finland	2.1815	2.1846	2.1782	2.1838	2.1810	2.1894	2.2007
France	9.2809	9.3937	9.5066	9.6157	9.7033	9.8065	9.9033
Germany	10.8188	10.6471	10.4725	10.2922	10.1228	9.9497	9.7842
Greece	2.8738	2.8540	2.8470	2.8318	2.8210	2.8118	2.7980
Hungary	2.6473	2.6202	2.6065	2.5846	2.5772	2.5598	2.5480
Ireland	2.2157	2.2406	2.2554	2.2817	2.2992	2.3274	2.3486
Italy	8.5522	8.5753	8.5994	8.6206	8.6085	8.5886	8.5371

(continued)

Table 2 (continued)

Country	2030	2035	2040	2045	2050	2055	2060
Latvia	1.7697	1.7800	1.7800	1.7911	1.7930	1.7941	1.8079
Lithuania	1.8941	1.8922	1.8909	1.8897	1.8905	1.8910	1.8933
Luxembourg	1.6087	1.6216	1.6464	1.6596	1.6738	1.6870	1.7014
Malta	1.5857	1.5989	1.6131	1.6267	1.6412	1.6546	1.6694
Netherlands	3.4973	3.4849	3.4535	3.4269	3.3960	3.3680	3.3593
Poland	4.9053	4.7567	4.6147	4.4920	4.3952	4.3233	4.2818
Portugal	2.8286	2.8316	2.8362	2.8318	2.8210	2.8222	2.8188
Romania	3.8166	3.7404	3.6695	3.5953	3.5189	3.4590	3.3893
Slovakia	2.1475	2.1400	2.1341	2.1291	2.1273	2.1146	2.1054
Slovenia	1.7697	1.7800	1.7910	1.7911	1.8036	1.8047	1.8186
Spain	7.4238	7.4884	7.5629	7.6325	7.6754	7.6981	7.6707
Sweden	2.7154	2.7203	2.7270	2.7352	2.747	2.7596	2.7774
UK	9.4323	9.6052	9.7642	9.9426	10.1228	10.3042	10.4791
Total	100.0002	99.9997	100.0002	99.9999	99.9999	100.0001	100.0001

Finally, we also present for each of the 13 years under consideration the a priori betting odds against an act being approved by the CM. These odds are just a modified form of A.

Note that A, R and the betting odds should not be interpreted too literally. Clearly, the CM does not vote on acts at random. Before an act is tabled for a formal vote at the CM, it goes through a preparatory process of bargaining and successive modification, until a point is reached where its approval is normally a foregone conclusion. What A, R and the betting odds actually measure is the average ease or difficulty of the preparatory process and the brevity or length of the time it may be expected to take.

3 Presentation of Results

The results of our calculations, presented in the Appendix, are organized as follows.

All our results are in the form of 13-term time series, consisting of data for the years 2008, 2011, 2014, and then for 2015–2060 at five-year intervals: 2015, 2020, ..., 2060. The values for 2008, 2011 and 2014 are calculated under the Nice QM rule; those from 2015 on are calculated under the QM rule of the Lisbon Treaty.

Tables 1, 2, 3 and 4 present, for each member-state and date, the respective values of four quantities: ψ (psi), β (beta), γ (gamma) and Quotient. The meaning of these quantities is the same as in our previous papers (Felsenthal and Machover 2001, 2004a, b, 2007). We recapitulated their explanation in Sect. 2.

Table 5 presents a synoptic comparison of various global properties—equitability, conformity to majority rule, sensitivity, efficiency—of the decision rules

Table 3 γ values by country and year

Country	2008	2011	2014	2015	2020	2025
Austria	0.32060	0.32060	0.32060	0.17047	0.17073	0.17190
Belgium	0.38196	0.38196	0.38196	0.19024	0.19100	0.19185
Bulgaria	0.32060	0.32060	0.32060	0.16087	0.15886	0.15720
Cyprus	0.12961	0.12961	0.12961	0.10905	0.10997	0.11065
Czech Rep.	0.38196	0.38196	0.38196	0.18549	0.18474	0.18416
Denmark	0.22610	0.22610	0.22610	0.14660	0.14708	0.14716
Estonia	0.12961	0.12961	0.12961	0.11224	0.11235	0.11299
Finland	0.22610	0.22610	0.22610	0.14501	0.14474	0.14561
France	0.80686	0.80686	0.80686	0.60209	0.60826	0.61276
Germany	0.80687	0.80687	0.80687	0.75790	0.74650	0.73402
Greece	0.38196	0.38196	0.38196	0.19341	0.19334	0.19262
Hungary	0.38196	0.38196	0.38196	0.18153	0.18004	0.17877
Ireland	0.22610	0.22610	0.22610	0.14262	0.14474	0.14716
Italy	0.80686	0.80686	0.80686	0.57411	0.57323	0.57115
Latvia	0.12961	0.12961	0.12961	0.11944	0.11946	0.11923
Lithuania	0.22610	0.22610	0.22610	0.12831	0.12739	0.12765
Luxembourg	0.12961	0.12961	0.12961	0.10583	0.10680	0.10750
Malta	0.09768	0.09768	0.09768	0.10502	0.10521	0.10597
Netherlands	0.41199	0.41199	0.41238	0.23492	0.23511	0.23512
Poland	0.76923	0.76923	0.76923	0.35860	0.34840	0.33870
Portugal	0.38196	0.38196	0.38196	0.18865	0.18943	0.18954
Romania	0.44155	0.44155	0.44155	0.27095	0.26643	0.26176
Slovakia	0.22610	0.22610	0.22610	0.14501	0.14552	0.14484
Slovenia	0.12961	0.12961	0.12961	0.11865	0.11867	0.11845
Spain	0.76925	0.76925	0.76925	0.47587	0.48613	0.49090
Sweden	0.32060	0.32060	0.32060	0.17836	0.18004	0.18109
UK	0.80686	0.80686	0.80686	0.59862	0.60912	0.61874

Country	2030	2035	2040	2045	2050	2055	2060
Austria	0.17207	0.17244	0.17214	0.17199	0.17199	0.17203	0.17164
Belgium	0.19181	0.19331	0.19335	0.19352	0.19390	0.19367	0.19444
Bulgaria	0.15533	0.15301	0.15158	0.15028	0.14914	0.14806	0.14715
Cyprus	0.11165	0.11236	0.11384	0.11456	0.11604	0.11672	0.11747
Czech Rep.	0.18270	0.18138	0.18019	0.17845	0.17767	0.17624	0.17513
Denmark	0.14769	0.14851	0.14862	0.14883	0.14914	0.14946	0.14996
Estonia	0.11320	0.11312	0.11384	0.11456	0.11532	0.11601	0.11606
Finland	0.14617	0.14626	0.14568	0.14592	0.14555	0.14593	0.14644
France	0.62187	0.62890	0.63584	0.64252	0.64757	0.65363	0.65899
Germany	0.72492	0.71282	0.70045	0.68772	0.67556	0.66318	0.65106
Greece	0.19256	0.19107	0.19042	0.18922	0.18827	0.18741	0.18619
Hungary	0.17738	0.17542	0.17434	0.17270	0.17199	0.17062	0.16955
Ireland	0.14846	0.15001	0.15085	0.15246	0.15344	0.15513	0.15628

(continued)

Table 3 (continued)

Country	2030	2035	2040	2045	2050	2055	2060
Italy	0.57305	0.57411	0.57516	0.57602	0.57450	0.57245	0.56807
Latvia	0.11858	0.11917	0.11905	0.11968	0.11966	0.11958	0.12030
Lithuania	0.12691	0.12668	0.12647	0.12627	0.12616	0.12604	0.12599
Luxembourg	0.10780	0.10856	0.11012	0.11090	0.11171	0.11244	0.11321
Malta	0.10625	0.10705	0.10789	0.10870	0.10952	0.11028	0.11109
Netherlands	0.23434	0.23331	0.23098	0.22898	0.22663	0.22449	0.22354
Poland	0.32868	0.31846	0.30865	0.30015	0.29332	0.28816	0.28492
Portugal	0.18953	0.18958	0.18970	0.18922	0.18827	0.18811	0.18757
Romania	0.25574	0.25042	0.24543	0.24023	0.23484	0.23056	0.22553
Slovakia	0.14390	0.14327	0.14274	0.14227	0.14197	0.14094	0.14010
Slovenia	0.11858	0.11917	0.11979	0.11968	0.12037	0.12029	0.12101
Spain	0.49744	0.50134	0.50584	0.51000	0.51223	0.51310	0.51042
Sweden	0.18194	0.18212	0.18239	0.18277	0.18332	0.18394	0.18481
UK	0.63202	0.64306	0.65307	0.66436	0.67556	0.68680	0.69730

operating at each of the 13 dates. For a brief explanation of the parameters used for this comparison, see Sect. 2.

Table 6 presents the EUROSTAT population data and forecasts. This table is copied from http://tinyurl.com/6kj56m.

Finally, Table 7—derived directly from Table 6—gives the rank-order of the member-states according to population size for each of the 13 dates of the latter table.

4 Analysis of the Results

First let us address the changes between 2008 and 2015. These are essentially the same as those described in our report (Felsenthal and Machover 2007) in which we compared the Nice rule with the rule that is now incorporated in the Lisbon Treaty. Although in Felsenthal and Machover (2007) we assumed the 2006 population data for both rules—rather than the 2008 forecast for the former and the 2015 forecast for the latter—the overall picture is the same. Let us summarize these changes.

Our projections show that all member-states will have in 2015 under the Lisbon Treaty rule more absolute voting power (as measured by ψ) than in 2008 under the Nice rule, but the increase is very uneven, not to say erratic.

The *relative* position (as measured by β) of the four largest (France, Germany, Italy, UK) and six smallest (Cyprus, Estonia, Latvia, Luxembourg, Malta, Slovenia) member-states will improve considerably, and that of Denmark will improve very slightly. The relative position of all other member-states will be worsened; the greatest loss of relative power will be sustained by Poland, followed

Table 4 Quotient values by country and year

Country	2008	2011	2014	2015	2020	2025
Austria	1.0306	1.0299	1.0289	0.8462	0.8443	0.8468
Belgium	1.0859	1.0824	1.0791	0.8309	0.8291	0.8282
Bulgaria	1.0762	1.0905	1.1037	0.8603	0.8655	0.8732
Cyprus	1.3493	1.3245	1.2996	1.6815	1.6441	1.6098
Czech Rep.	1.1020	1.1045	1.1064	0.8319	0.8310	0.8330
Denmark	0.8966	0.8973	0.8978	0.9009	0.9029	0.9013
Estonia	1.0396	1.0485	1.0558	1.4178	1.4332	1.4584
Finland	0.9114	0.9117	0.9114	0.9045	0.9071	0.9067
France	0.9519	0.9493	0.9465	1.0919	1.0969	1.0994
Germany	0.8260	0.8310	0.8359	1.2173	1.2080	1.1970
Greece	1.0584	1.0582	1.0585	0.8296	0.8307	0.8305
Hungary	1.1184	1.1266	1.1340	0.8356	0.8361	0.8381
Ireland	0.9986	0.9724	0.9509	0.9221	0.9094	0.9063
Italy	0.9705	0.9702	0.9704	1.0688	1.0683	1.0667
Latvia	0.7985	0.8087	0.8177	1.1701	1.1896	1.2083
Lithuania	1.1437	1.1573	1.1694	1.0302	1.0370	1.0537
Luxembourg	1.7321	1.7105	1.6897	2.1263	2.1014	2.0725
Malta	1.4149	1.4150	1.4142	2.3521	2.3515	2.3679
Netherlands	0.9440	0.9452	0.9468	0.8349	0.8355	0.8348
Poland	1.1563	1.1633	1.1690	0.8446	0.8260	0.8101
Portugal	1.0878	1.0860	1.0845	0.8285	0.8301	0.8299
Romania	0.8853	0.8931	0.9003	0.8571	0.8526	0.8484
Slovakia	0.9030	0.9070	0.9104	0.9044	0.9063	0.9141
Slovenia	0.8457	0.8471	0.8488	1.2032	1.2082	1.2145
Spain	1.0608	1.0438	1.0313	0.9840	0.9932	0.9976
Sweden	0.9818	0.9777	0.9737	0.8370	0.8378	0.8361
UK	0.9566	0.9537	0.9500	1.0891	1.0978	1.1044

Country	2030	2035	2040	2045	2050	2055	2060
Austria	0.8427	0.8411	0.8371	0.8346	0.8331	0.8321	0.8290
Belgium	0.8278	0.8232	0.8187	0.8152	0.8126	0.8070	0.8052
Bulgaria	0.8776	0.8794	0.8850	0.8904	0.8968	0.9035	0.9123
Cyprus	1.5834	1.5590	1.5478	1.5272	1.5180	1.4996	1.4845
Czech Rep.	0.8310	0.8309	0.8304	0.8263	0.8267	0.8243	0.8244
Denmark	0.8998	0.9016	0.9001	0.8995	0.8989	0.8971	0.8949
Estonia	1.4763	1.4911	1.5130	1.5329	1.5526	1.5714	1.5837
Finland	0.9094	0.9116	0.9107	0.9142	0.9125	0.9139	0.9149
France	1.1074	1.1123	1.1171	1.1220	1.1243	1.1277	1.1292
Germany	1.1889	1.1773	1.1663	1.1552	1.1455	1.1348	1.1238
Greece	0.8250	0.8252	0.8224	0.8174	0.8144	0.8130	0.8108
Hungary	0.8383	0.8362	0.8373	0.8346	0.8362	0.8341	0.8339
Ireland	0.8988	0.8956	0.8884	0.8853	0.8787	0.8770	0.8733

(continued)

Table 4 (continued)

Country	2030	2035	2040	2045	2050	2055	2060
Italy	1.0697	1.0714	1.0729	1.0750	1.0743	1.0740	1.0703
Latvia	1.2212	1.2476	1.2644	1.2877	1.3039	1.3198	1.3467
Lithuania	1.0613	1.0750	1.0886	1.1019	1.1160	1.1301	1.1461
Luxembourg	2.0320	2.0050	1.9958	1.9754	1.9578	1.9393	1.9218
Malta	2.3746	2.4018	2.4327	2.4616	2.4887	2.5115	2.5354
Netherlands	0.8294	0.8249	0.8175	0.8126	0.8065	0.8001	0.7967
Poland	0.7936	0.7783	0.7639	0.7522	0.7441	0.7400	0.7414
Portugal	0.8272	0.8252	0.8234	0.8194	0.8143	0.8134	0.8115
Romania	0.8386	0.8307	0.8236	0.8153	0.8067	0.8018	0.7961
Slovakia	0.9150	0.9204	0.9270	0.9339	0.9425	0.9468	0.9540
Slovenia	1.2241	1.2405	1.2574	1.2667	1.2854	1.2966	1.3176
Spain	1.0065	1.0116	1.0178	1.0236	1.0275	1.0307	1.0286
Sweden	0.8336	0.8305	0.8279	0.8248	0.8212	0.8169	0.8137
UK	1.1153	1.1238	1.1304	1.1384	1.1453	1.1515	1.1563

Table 5 Synoptic comparison

Year	D	max|d|	ran(d)	MMD	S	A	R	Odds
2008	4.7118	73.2	93.4	8037	0.858	0.020	0.959	49:1
2011	4.6931	71.0	90.2	8084	0.858	0.020	0.959	49:1
2014	4.6813	69.0	87.2	8125	0.858	0.020	0.959	49:1
2015	7.5633	135.2	152.3	5368	0.945	0.128	0.745	34:5
2020	7.6182	135.2	152.3	5040	0.945	0.127	0.745	336:49
2025	7.6546	136.8	154.0	5413	0.945	0.128	0.744	34:5
2030	7.8467	137.5	158.1	5446	0.944	0.127	0.746	336:49
2035	8.0014	140.2	157.9	5452	0.944	0.127	0.746	336:49
2040	8.1769	143.3	161.4	5450	0.944	0.127	0.746	336:49
2045	8.3401	146.2	164.9	5446	0.944	0.127	0.746	336:49
2050	8.4675	148.9	167.6	5428	0.944	0.127	0.746	336:49
2055	8.5628	151.2	171.1	5406	0.944	0.127	0.746	336:49
2060	8.5938	153.5	173.0	5371	0.944	0.127	0.745	336:49

by Hungary and the Czech Republic. The smallest loss will be experienced by Finland and Slovakia.[11]

As for blocking power, γ, Malta will gain slightly; all other member-states will lose blocking power, but the extent of loss is again very uneven.

From Table 4 we can see that, by the yardstick of Penrose's Square-Root Rule, the voting-power distribution in 2015 will be considerably less equitable than in 2008. As can be seen from this table under the 2015 column, the two most

[11] In fact Denmark, as well as Finland and Slovakia, will experience the smallest change in β, and consequently in their equitability Quotient.

Table 6 Population forecast of present EU members (1000s)

Country	2008	2011	2014	2015	2020	2025
Austria	8,334	8,439	8,538	8,570	8,723	8,866
Belgium	10,656	10,844	11,016	11,070	11,322	11,547
Bulgaria	7,642	7,527	7,419	7,382	7,188	6,974
Cyprus	795	834	874	888	955	1,017
Czech Rep.	10,346	10,417	10,480	10,497	10,543	10,516
Denmark	5,476	5,530	5,577	5,591	5,661	5,736
Estonia	1,339	1,331	1,325	1,323	1,311	1,292
Finland	5,300	5,356	5,412	5,429	5,501	5,549
France	61,876	62,921	63,896	64,203	65,607	66,846
Germany	82,179	82,098	81,919	81,858	81,472	80,907
Greece	11,217	11,347	11,450	11,476	11,556	11,575
Hungary	10,045	10,011	9,976	9,964	9,893	9,790
Ireland	4,415	4,709	4,971	5,052	5,404	5,673
Italy	59,529	60,233	60,784	60,929	61,421	61,683
Latvia	2,269	2,237	2,209	2,200	2,151	2,095
Lithuania	3,365	3,324	3,287	3,275	3,220	3,158
Luxembourg	482	500	517	523	551	579
Malta	410	415	420	421	427	431
Netherlands	16,404	16,548	16,679	16,717	16,896	17,069
Poland	38,116	38,080	38,073	38,068	37,960	37,612
Portugal	10,617	10,773	10,908	10,947	11,108	11,224
Romania	21,423	21,287	21,148	21,103	20,834	20,484
Slovakia	5,399	5,411	5,423	5,427	5,432	5,402
Slovenia	2,023	2,039	2,050	2,053	2,058	2,047
Spain	45,283	47,301	48,924	49,381	51,109	52,101
Sweden	9,183	9,364	9,533	9,588	9,853	10,094
UK	61,270	62,340	63,424	63,792	65,683	67,543
Total	495,393	501,216	506,232	507,727	513,839	517,810

Country	2030	2035	2040	2045	2050	2055	2060
Austria	8,988	9,075	9,122	9,138	9,127	9,088	9,037
Belgium	11,745	11,906	12,033	12,125	12,194	12,247	12,295
Bulgaria	6,753	6,535	6,330	6,129	5,923	5,710	5,485
Cyprus	1,072	1,121	1,167	1,211	1,251	1,288	1,320
Czech Rep.	10,420	10,288	10,158	10,036	9,892	9,722	9,514
Denmark	5,808	5,858	5,882	5,890	5,895	5,903	5,920
Estonia	1,267	1,243	1,221	1,202	1,181	1,159	1,132
Finland	5,569	5,557	5,521	5,481	5,448	5,422	5,402
France	67,982	69,021	69,898	70,553	71,044	71,442	71,800
Germany	80,152	79,150	77,821	76,249	74,491	72,621	70,759
Greece	11,573	11,575	11,567	11,531	11,445	11,301	11,118
Hungary	9,651	9,501	9,352	9,213	9,061	8,898	8,717

(continued)

Table 6 (continued)

Country	2030	2035	2040	2045	2050	2055	2060
Ireland	5,881	6,057	6,221	6,381	6,531	6,654	6,752
Italy	61,868	61,995	62,002	61,777	61,240	60,413	59,390
Latvia	2,033	1,970	1,913	1,858	1,804	1,746	1,682
Lithuania	3,083	2,998	2,912	2,825	2,737	2,645	2,548
Luxembourg	607	633	657	678	697	715	732
Malta	432	429	424	419	415	410	405
Netherlands	17,208	17,271	17,226	17,085	16,909	16,740	16,596
Poland	36,975	36,141	35,219	34,257	33,275	32,244	31,139
Portugal	11,317	11,395	11,452	11,475	11,449	11,373	11,265
Romania	20,049	19,619	19,161	18,679	18,149	17,584	16,921
Slovakia	5,332	5,231	5,115	4,993	4,859	4,712	4,547
Slovenia	2,023	1,992	1,958	1,921	1,878	1,830	1,779
Spain	52,661	53,027	53,290	53,409	53,229	52,701	51,913
Sweden	10,270	10,382	10,470	10,565	10,672	10,780	10,875
UK	69,224	70,685	72,009	73,282	74,506	75,647	76,677
Total	519,943	520,655	520,101	518,362	515,302	510,995	505,720

egregious cases are: on the one hand Malta, which will have 135.2 % more than its fair share; and on the other hand Portugal, which will have 17.15 % too little.

From Table 5 we observe that the Lisbon Treaty QM rule is quite efficient: it has a relatively high value of Coleman's index *A* (the a priori probability of approving an act rather than blocking it) and a correspondingly low resistance R. In betting terms, this means that the a priori odds against approval of an act will be approximately 34 to 5 in 2015 and subsequently 336 to 49. This is a very considerable improvement compared to the Nice rule, which is extremely (and dangerously) inefficient.

With respect to sensitivity (S) and mean majority deficit (MMD), the Lisbon Treaty QM rule is also a definite improvement compared to the Nice rule.

Now let us turn to the period 2015–2060. As can be seen in Table 6, according to EUROSTAT forecasts the total EU population will continue to grow until 2035, reaching its maximal size of 520.6 million. Beginning in 2040, the total EU population decreases gradually, reaching its smallest size of 505.7 million in 2060, with the steepest drop of 5.27 million occurring between 2055 and 2060. But different groups of countries will undergo quite distinct demographic changes.

The populations of all ten Eastern European and Baltic EU member-states decrease steadily throughout the period 2008–2060. The steepest decrease (in both absolute and relative terms) among these countries is experienced by Poland which is expected to lose 6.9 million people (18.3 %) between 2008 and 2060.

Among the remaining EU member-states, relatively significant decreases in populations are expected to occur in Germany and Italy, while relatively significant increases are expected to occur in France and the United Kingdom.

As a result of these different population changes in individual countries, the rank-order according to population size of only nine EU members (Finland, Italy,

Table 7 Country by population rank by year

Country	2008	2011	2014	2015	2020	2025
Austria	15	15	15	15	15	15
Belgium	10	10	10	10	10	10
Bulgaria	16	16	16	16	16	16
Cyprus	25	25	25	25	25	25
Czech Rep.	12	12	12	12	12	12
Denmark	17	17	17	17	17	17
Estonia	24	24	24	24	24	24
Finland	19	19	19	19	19	19
France	2	2	2	2	3	3
Germany	1	1	1	1	1	1
Greece	9	9	9	9	9	9
Hungary	13	13	13	13	13	14
Ireland	20	20	20	20	20	18
Italy	4	4	4	4	4	4
Latvia	22	22	22	22	22	22
Lithuania	21	21	21	21	21	21
Luxembourg	26	26	26	26	26	26
Malta	27	27	27	27	27	27
Netherlands	8	8	8	8	8	8
Poland	6	6	6	6	6	6
Portugal	11	11	11	11	11	11
Romania	7	7	7	7	7	7
Slovakia	18	18	18	18	18	20
Slovenia	23	23	23	23	23	23
Spain	5	5	5	5	5	5
Sweden	14	14	14	14	14	13
UK	3	3	3	3	2	2

Country	2030	2035	2040	2045	2050	2055	2060
Austria	15	15	15	15	14	14	14
Belgium	10	9	9	9	9	9	9
Bulgaria	16	16	16	17	17	18	18
Cyprus	25	25	25	24	24	24	24
Czech Rep.	12	13	13	13	13	13	13
Denmark	18	18	18	18	18	17	17
Estonia	24	24	24	25	25	25	25
Finland	19	19	19	19	19	19	19
France	3	3	3	3	3	3	2
Germany	1	1	1	1	2	2	3
Greece	9	10	10	10	11	11	11
Hungary	14	14	14	14	15	15	15
Ireland	17	17	17	16	16	16	16

(continued)

Table 7 (continued)

Country	2030	2035	2040	2045	2050	2055	2060
Italy	4	4	4	4	4	4	4
Latvia	22	23	23	23	23	23	23
Lithuania	21	21	21	21	21	21	21
Luxembourg	26	26	26	26	26	26	26
Malta	27	27	27	27	27	27	27
Netherlands	8	8	8	8	8	8	8
Poland	6	6	6	6	6	6	6
Portugal	11	11	11	11	10	10	10
Romania	7	7	7	7	7	7	7
Slovakia	20	20	20	20	20	20	20
Slovenia	23	22	22	22	22	22	22
Spain	5	5	5	5	5	5	5
Sweden	13	12	12	12	12	12	12
UK	2	2	2	2	1	1	1

Lithuania, Luxembourg, Malta, Netherlands, Poland, Romania, and Spain) will remain unchanged during the entire period 2008–2060. The rank-order of the remaining 18 EU members is expected to change at least once during this period (cf. Table 7).

The ψ, β and γ values for the various member-states during the period 2015–2060 are of course consistent with both the absolute and relative sizes of their respective populations. Thus, for example, the values of these three measures for Malta are smaller than those of any other EU member in any given period because Malta's population size ranks 27 in all periods. However, the changes from one period to the next are quite small.

Finally, as can be observed from Table 5, the changes from one period to the next during 2015–2060 of each of the synoptic parameters are very small and insignificant.

4.1 Conclusions

Not surprisingly, our computations show that the main changes in voting power and related quantities will occur in the change-over from the Nice QM rule to the QM rule of the Lisbon Treaty, which in our projection will have taken place between 2008 and 2015. From 2015 on the changes—due entirely to demographic trends—are relatively small.

Acknowledgments Voting Power and Procedures Programme: Supported by Voting Power in Practice Grant F/07 004/AJ from the Leverhulme Trust, which the authors gratefully acknowledge.

References

11177/1/07 REV 1. (2007). *Council of the European Union: Brussels European Council 21/22 June 2007—Presidency Conclusions, Brussels*, 20 July 2007. EU Document 11177/1/07 REV 1.

CIG 85/04. (2004). *Conference of the Representatives of the Governments of the Member States*, Brussels, 18 June 2004. EU document CIG 85/04.

Eurostat (the statistical office of the European Commission), project EUROPOP2008 (EUROpean POpulation Projections, base year 2008).

Felsenthal, D. S., & Machover, M. (1998). *The measurement of voting power: Theory and practice, problems and paradoxes*. Cheltenham: Edward Elgar.

Felsenthal, D. S., & Machover, M. (2001). The Treaty of Nice and qualified majority voting. *Social Choice and Welfare*, *18*(3), 431–464. http://eprints.lse.ac.uk/archive/420/.

Felsenthal, D. S., & Machover, M. (2004a). Analysis of QM rules in the draft Constitution for Europe proposed by the European Convention, 2003. *Social Choice and Welfare 23*(1), 1–20. http://eprints.lse.ac.uk/archive/429/.

Felsenthal, D. S., & Machover, M. (2004b). *Analysis of QM rule adopted by the EU Inter-Governmental Conference Brussels, 18 June 2004, Project Report*. London, UK: London School of Economics and Political Science. http://eprints.lse.ac.uk/archive/431/

Felsenthal, D. S., & Machover, M. (2007). *Analysis of QM Rule adopted by the Council of the European Union, Brussels, 23 June 2007, Project Report*. London, UK: London School of Economics and Political Science. http://eprints.lse.ac.uk/archive/2531/.

Nurmi, H. (2008). *On the relative unimportance of voting weights. Observations on Agenda-Based Voting Procedures*. Mimeograph. http://tinyurl.com/ad7yt6

Treaty of Lisbon. (2007). Treaty of Lisbon amending the Treaty on European Union and the Treaty establishing the European Community, signed at Lisbon, 13 December 2007. *Official Journal of the European Union 50*, notice no. 2007/C306/ 17 December 2007.

Treaty of Nice. (2001). Conference of the Representatives of the Governments of the Member States, Brussels, 28 February 2001. *Treaty of Nice amending the Treaty on European Union, the Treaties establishing the European Communities and certain related Acts*. EU document confer 4820/00.

Widgrén, M. (2008). *The impact of Council's internal decision-making rules on the future EU*. Mimeograph. CESifo working paper no. 2195. http://tinyurl.com/c8lxur.

Part VI
The Aggregation of Preferences

Explaining All Possible Paired Comparison Problems

Donald G. Saari

1 Introduction

Already in the eighteenth century, Condorcet (1785) recognized the negative effects of voting cycles when he promoted the use of majority votes of pairs. Beyond voting theory, cycles and related anomalies have created problems affecting a variety of disciplines when paired comparisons are being used. Whether utilizing this approach to make engineering choices, analyze psychological experiments, or examine economic theory, cycles can frustrate the outcomes by introducing ambiguous outcomes.

The same difficulty arises in statistics where cyclic behavior has been analyzed for decades, while in probability theory independence conditions are imposed to contain these concerns. Cycles, in other words, create problems that cut across disciplines. As such, an explanation why these complications arise and how to avoid them would be welcomed. Such a structure exists; the source of all possible troubles, paradoxes, and anomalies caused by examining pairs can be explained with what I call a "ranking wheel" (Saari 2008).

A related concern is to understand why outcomes can differ when the information (e.g., a profile of voter preferences) about several alternatives is analyzed in terms of pairs or in terms of the 'whole'. In economics, for instance, the difference between what happens with bilateral trades among three countries and an arrangement among all three can be significant. Differences between the parts and the whole also arise with election rankings. As an example, the Borda Count (tally a N-candidate ballot by assigning $N - j$ points to the jth ranked candidate) has been criticized because it need not elect a Condorcet winner (i.e., a candidate who beats all others in pairwise majority vote elections); a long standing concern has been to understand why this can happen. The Analytic Hierarchy Process

D. G. Saari (✉)
Institute for Mathematical Behavioral Sciences,
University of California, Irvine, CA 92697-5100, USA
e-mail: dsaari@uci.edu

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*, DOI: 10.1007/978-3-642-35929-3_32,

(AHP) (Saaty (1980); described below) also ranks options. A new PhD blessed with several job offers might use AHP to determine which university position, A, B, or C, to accept based on how she evaluates the attributes of these schools. One of several problems with AHP, however, is that its three-option ranking can disagree with how two options are ranked with the same data; e.g., the $A \succ B \succ C$ ranking could be contradicted by a $C \succ A$ ranking. Which university should our candidate select; A or C? Rather than the candidate's decision, our concern is to understand what causes this AHP behavior.

More generally, the seminal Arrow's Theorem (1951) suggests that a group outcome for three candidates (or alternatives) need not agree with the paired comparison outcomes reached by the same voters. A long standing mystery is to understand what causes this conclusion. A similar issue surfaces in Sen's result (1970a, b) demonstrating a conflict between where certain individuals can make personal decisions and where, for other pairs, the unanimous consent of a ranking determines that pair's societal ranking. How can this be true! Could a more comforting explanation be found that differs from Sen's?

A variety of further issues range from the discursive paradox in philosophy (recently discussed by Christian List and coauthors) to the mysterious Nakamura's number (1978) identifying the maximum number of alternatives with which cycles cannot occur. To introduce the spirit of the first comment, suppose a panel of three is evaluating whether an Assistant Professor should receive tenure; the decision is made by a majority vote. Each panel member is instructed to ensure that the candidate excels in both research and teaching. Unfortunately, the candidate failed with the following scorecard:

Panel member	Research	Teaching	Decision
A	No	Yes	No
B	Yes	No	No
C	Yes	Yes	Yes
Outcome			**No**

(1)

But notice; by majority 2:1 votes, the panel finds that this candidate's performance in teaching and in research is satisfactory. Had this panel voted over each criterion to make its decision, our candidate would have been promoted rather than forced to search for a new position. What causes this difficulty? (A "Yes-No" vote is a paired comparison.)

The Nakamura number addresses the reality that not all paired comparisons involve a simple majority vote. To avoid a filibuster in the United States Senate, for instance, a bill must receive a supermajority of 60 of the 100 votes. More generally, a q-rule is where a quota q of the n votes (where $q > \frac{n}{2}$) are needed to pass a proposition. By imposing a more stringent condition for success, q-rules make it more difficult to have cycles. But they can arise; Nakamura determined the minimum number of candidates with which a q-rule cycle becomes possible. As an illustration, if victory in a paired comparison requires 72 of the 84 possible votes, it takes at least seven alternatives to create a cycle. This means that any setting with six or fewer alternatives is spared a cycle with this q-rule. Why?

One must wonder whether a common structure can be found to provide insight into these seemingly varied issues; e.g., is there a simple, common way to understand Arrow's, Sen's, and the other described assertions? Similarly, a flawed but standard way to analyze decision rules is in terms of whether they do, or do not, satisfy certain properties; can we go beyond the severe limitations of this (incorrectly) called "axiomatic approach" to understand *why* certain properties are, or are not, satisfied? The answer for paired comparisons is yes; answers follow from the "ranking wheel."

Beyond addressing all of these kinds of difficulties, there are reasons to believe that this ranking wheel approach explains, or helps to explain, all possible paired comparison problems that come from any discipline. This structure even partly explains other types of difficulties that arise with triplets. But as explanations for triplets, and higher order clusterings, require introducing other kinds of ranking wheels, this material will be described elsewhere. (For $N = 3$, see Saari 2008, Chap. 4 for a complete description.)

2 Examples

Examples are given to illustrate two of the above comments.

2.1 A Voting Example

Suppose 33 voters select an alternatives from $\{A, B, C\}$ where their preferences are

Number	Ranking	Number	Ranking
6	$A \succ B \succ C$	5	$C \succ B \succ A$
9	$A \succ C \succ B$	2	$B \succ C \succ A$
1	$C \succ A \succ B$	10	$B \succ A \succ C$

(2)

The pairwise majority vote outcomes are $B \succ A$ and $B \succ C$ by respective tallies of 17:16 and 18:15, so B is the Condorcet winner. (In the final pair, $A \succ C$ by 25:8.) The Borda tally (where 2, 1, and 0 points are assigned, respectively, to a ballot's top, second, and bottom ranked candidates) yields the conflicting $A \succ B \succ C$ ranking (with tally 41:35:23) that *demotes the Condorcet winner B to second place!* A complete understanding of this phenomenon comes from the ranking wheel.

2.2 An AHP Example

The intent of AHP is to go beyond ranking the alternatives to determine an intensity value assigned to each alternative. With alternatives (or options) A_i and A_j, the $a_{i,j}$ value is defined to be the multiple of how much an evaluator believes that A_i is better than A_j. With these scale multiples, it follows that $a_{j,i} = 1/a_{i,j}$ and that $a_{j,j} = 1$. These values define a matrix $\mathcal{M}_N$; e.g., a simple example is

$$\mathcal{M}_3 = \begin{pmatrix} 1 & 6 & \frac{3}{4} \\ \frac{1}{6} & 1 & 8 \\ \frac{4}{3} & \frac{1}{8} & 1 \end{pmatrix} \tag{3}$$

where, for instance, $a_{1,2} = 6$ means that the evaluator views A_1 as being six times better than A_2, while $a_{2,1} = \frac{1}{a_{1,2}} = \frac{1}{6}$ means that A_2 is one-sixth as good as A_1.

Because $\mathcal{M}_N$ has only positive entries, the Perron-Frobenius Theorem ensures that $\mathcal{M}_N$ has a unique eigenvector of positive terms, $\mathbf{w} = (w_1, w_2, , w_N)$. By normalizing these values so that $\sum w_j = 1$, component w_j defines A_j's portion of the whole. (The Eq. 3 normalized eigenvector is $\mathbf{w} = \left(\frac{3}{6}, \frac{2}{6}, \frac{1}{6}\right)$, which defines the $A_1 \succ A_2 \succ A_3$ ranking.) In this manner, AHP goes beyond ranking the alternatives to emphasize the w_j value, which is intended to define the "intensity" that the decision maker should assign to option A_j. Size matters where "larger is better." Illustrating with Eq. 3, options A_1, A_2, A_3 have, respectively, the weights of $\frac{1}{2}$, $\frac{1}{3}$, and $\frac{1}{6}$, so the intensity assigned to A_1 is three times that assigned to A_3. This multiple indicates the significant superiority of A_1 over A_3.

A way to justify using this eigenvector is in terms of an iterative "updating" process. Let $v_j > 0$ be an initial guess of the intensity assigned to option A_j where $\sum v_j = 1$. These estimates define a vector $\mathbf{v}_0 = (v_1, v_2, \ldots, v_N)$, which can be treated as a "prior." Following the spirit of Bayesian updating, refine $\mathbf{v}_0$ by using the paired comparison information catalogued in the matrix $\mathcal{M}_N$. The first row of $\mathcal{M}_N$, for instance, compares how A_1 fares with all other options. The product $a_{1,2}v_2$ captures how much better A_1 is than A_2 (given by $a_{1,2}$) times the influence of A_2 (given by v_2), so it provides a measure of the relative "influence" standing of the two options. To compare A_1 with all N options, the scalar product of this row with $\mathbf{v}_0$ (given by $\sum a_{1,j}v_j$) upgrades the intensity value for option A_1. Doing so for all options leads to $\mathbf{v}_1^t = \mathcal{M}_N(\mathbf{v}_0^t)$ (where superscript 't' indicates a transpose to make it a column vector). (Illustrating with Eq. 3 and a noncommittal $\mathbf{v}_0 = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$, the updated $\mathbf{v}_1$ is $\left(\frac{186}{465}, \frac{220}{465}, \frac{59}{465}\right)$.) By continuing to iterate, one might hope that the resulting, normalized $\mathbf{v}_n$ vectors converge to a limiting conclusion. They do; they converge to the (normalized) eigenvector $\mathbf{w}$.

To compare the extremes of the "best" and "worst" A_1 and A_3 options, information about A_2 is irrelevant, so ignore it. (That is, drop the second row and second column from $\mathcal{M}_3$; what remains compares A_1 and A_3.) The A_1 and A_3 information creates the matrix

$$\mathcal{M}_2 = \begin{pmatrix} 1 & \frac{3}{4} \\ \frac{4}{3} & 1 \end{pmatrix} \tag{4}$$

with eigenvector $\mathbf{w}^* = \left(\frac{3}{7}, \frac{4}{7}\right)$. Notice what has happened; with the same data, the three-option AHP outcome is $A_1 \succ A_2 \succ A_3$, but a direct comparison of the extremes creates the conflict by *reversing* the $\{A_1, A_3\}$ AHP-ranking to $A_3 \succ A_1$.

AHP's intensity weights are intended to capture "how much better" one option is than another. In the three-option setting, the weight assigned to A_1 is *three times* that assigned to A_3, which presumably indicates the significant superiority of A_1 over A_3. But directly comparing this pair using the same data yields the contradiction that A_3 is the dominant option with an intensity $\frac{4}{3}$ times greater than that of A_1! The challenge is to understand why this is so; the answer comes from the ranking wheel. (This wheel can be used to create examples where the A_1 intensity is any desired multiple larger than of A_3, but a direct comparison reverses the inequality. Similar examples can be created for any $N \geq 3$.)

3 The Ranking Wheel and a Coordinate System

To explain these difficulties, mount a freely rotating wheel (a disc) on a surface. With N candidates, place in an evenly spaced manner the numbers 1 through N along the edge of the wheel; these are the "ranking numbers." Next, select an initial ranking of the alternatives. On the surface, place each name next to the associated ranking number. Figure 1 depicts the $N = 6$ setting with the initial ranking $A \succ B \succ C \succ D \succ E \succ F$.

Rotate the ranking wheel so that ranking number 1 is by the name of the next alternative and write down the new ranking. In Fig. 1 , this is $B \succ C \succ D \succ E \succ F \succ A$. Continue to do this until the ranking number 1 has been next to each alternative once; this defines the "ranking wheel configuration" ($\mathcal{RWC}$) of N rankings. The Fig. 1 $\mathcal{RWC}$ is

$$\begin{array}{ll} A \succ B \succ C \succ D \succ E \succ F, & B \succ C \succ D \succ E \succ F \succ A, \\ C \succ D \succ E \succ F \succ A \succ B, & D \succ E \succ F \succ A \succ B \succ C, \\ E \succ F \succ A \succ B \succ C \succ D, & F \succ A \succ B \succ C \succ D \succ E. \end{array} \tag{5}$$

To illustrate with $N = 3$ and the starting $A \succ B \succ C$ ranking, the $\mathcal{RWC}$ defines the three rankings (called the "Condorcet triplet")

$$A \succ B \succ C, \quad B \succ C \succ A, \quad C \succ A \succ B. \tag{6}$$

With the reversed $C \succ B \succ A$ as the starting ranking, the other Condorcet triplet emerges:

$$C \succ B \succ A, \quad B \succ A \succ C, \quad A \succ C \succ B. \tag{7}$$

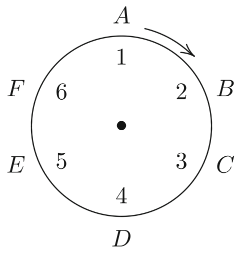


Fig. 1 The $N = 6$ ranking wheel

By construction a $\mathcal{RWC}$ lists each alternative in first, second, $, Nth$ position precisely once, so no alternative is favored over any other. But rather than a tie, the pairwise majority vote defines a cycle. With Eq. 6, this cycle is $A \succ B$, $B \succ C$, $C \succ A$ where each pair's tally is 2:1. A way to describe this cycle is that

> when C is not being considered, A wins; when A is not being considered, B wins; when B is not being considered, C wins.

The Eq. 5 profile creates the $A \succ B$, $B \succ C$, $C \succ D$, $D \succ E$, $E \succ F$, $F \succ A$ cycle where each pair has the more dramatic 5:1 tally. Indeed, cyclic behaviors generated by $\mathcal{RWC}$s have surprising tallies; e.g., a N-alternative $\mathcal{RWC}$ has a pairwise voting cycle (given by successive adjoining entries in the starting configuration) where each has a (N - 1):1 tally.

The construction makes it clear why cyclic rankings occur. Paired comparisons use only a minimal portion of the available information from a $\mathcal{RWC}$; e.g., the rule fails to recognize the global symmetry which demonstrates that the outcome should be a tie. Thus when a paired comparison rule evaluates a $\mathcal{RWC}$ profile, the cyclic construction of a $\mathcal{RWC}$ forces a cyclic action to emerge. (With any adjacent pair, such as A, B, the cyclic $\mathcal{RWC}$ construction keeps one of the alternatives ranked ahead of the other in all but one ranking. This cyclic construction, then, forces each alternative to dominate one alternative and be dominated by a different one.) This comment extends beyond paired comparisons to apply to any comparison involving "parts" (e.g., Saari and Sieberg 2004).

To illustrate a cyclic effect that involves other "parts," use the plurality vote to rank the six quintuplets defined by Eq. 5 where each quintuplet is determined by the "missing candidate." For example, if F is dropped, then A is the plurality winner of this six voter profile. Using the earlier wording, with these quintuplets,

> when F is not being considered, A is the plurality winner; when A is not being considered, B is the winner; when B is not being considered, C is the winner; when C is not being considered, D is the winner; when D is not being considered, E is the winner; and when E is not being considered, F is the winner,

to complete the quintuplet-wise cycle.

To introduce a different "part-whole" conflict, the paired tallies for Eq. 6 are

Ranking	$\{A, B\}$	$\{B, C\}$	$\{A, C\}$
$B \succ C \succ A$	$B \succ A$	$B \succ C$	$C \succ A$
$C \succ A \succ B$	$A \succ B$	$C \succ B$	$C \succ A$
$A \succ B \succ C$	$A \succ B$	$B \succ C$	$A \succ C$
Outcome	$A \succ B$	$B \succ C$	$C \succ A$

(8)

Notice how the 2:1 outcome for each pair agrees with that of the Eq. 1 example evaluating an Assistant Professor for tenure. To strengthen the similarities, replace each $A \succ B$, $B \succ C$, and $A \succ C$ in Eq. 8 with "Yes," and all other paired rankings with "No." What results is the Eq. 1 table.

This is no coincidence; all differences in conclusions caused by aggregating paired comparison information in different manners reflect ranking wheel structures. Thus, as with Eq. 8, expect it to be possible to create examples with $\mathcal{RWC}$s (or with Sect. 4 structures) that have paradoxical appearing conclusions. But also expect these conflicts to have similar explanations. With Eq. 6, the cycle occurs because the voting procedure forces global information—the transitivity of preferences—to essentially be ignored when computing outcomes for the parts. Similarly, if the more global information of each Eq. 1 evaluator's beliefs are to be valued, the negative outcome is appropriate. But if evaluators' total beliefs are secondary to what happens with "parts," tenure should be granted. Paired comparison issues of this type are captured by ranking wheel configurations.

A step toward understanding problems with paired comparisons is to determine the number of distinct trouble-causing $\mathcal{RWC}$s. With no restrictions on the choice of the initial ranking, each of the $N!$ rankings is in at least one $\mathcal{RWC}$. It is not difficult to prove the stronger statement that each ranking is in precisely one $\mathcal{RWC}$. But each $\mathcal{RWC}$ has N rankings, so there are $N!/N = (N-1)!$ different $\mathcal{RWC}$s. With $N = 6$, then, there are $5! = 120$ different $\mathcal{RWC}$s; with $N = 11$, there are $10! = 3,628,800$ $\mathcal{RWC}$s. What makes this rapid increase in the number of $\mathcal{RWC}$s of interest (and concern) is that they are totally responsible for all paired comparison problems.

3.1 Strongly Transitive

As $\mathcal{RWC}$ profiles cause all paired comparison problems, we might hope there exists an opposite extreme where "nothing goes wrong." Should such profiles exist, they can be viewed as being "anti-ranking wheel configurations." This comment suggests searching for these idealized settings by using the ranking wheel, but in a way to ensure that any emerging profile has nothing to do with any $\mathcal{RWC}$.

To illustrate how to do so, start with $N = 3$, which has has precisely two $\mathcal{RWC}$s given by Eqs. 6, 7. The sought after "well behaved" profiles strike a balance between these two configurations in that the number of voters with rankings coming from one $\mathcal{RWC}$ equals the number of voters with preference rankings coming from the other $\mathcal{RWC}$. To illustrate, all voters with a preference ranking coming from Eq. 6 are listed to the left in

Number	Ranking	Number	Ranking
6	$A \succ B \succ C$	4	$A \succ C \succ B$
1	$C \succ A \succ B$	0	$C \succ B \succ A$
2	$B \succ C \succ A$	5	$B \succ A \succ C$

(9)

while voters with preference rankings coming from Eq. 7 are placed to the right. With nine voters on each side of the line, this profile satisfies the balancing condition.

To interpret what has been gained, define the "tally difference," represented by $\tau(X, Y)$, to be the difference between the X and Y tallies when compared as a pair. Illustrating with Eq. 9 where the majority vote tallies are $A \succ B$ with 11:7, $B \succ C$ with 13:5, and $A \succ C$ with 15:3, the corresponding tally-difference values are

$$\tau(A, B) = 11 - 7 = 4, \ \ \tau(B, C) = 13 - 5 = 8, \textit{and} \ \ \tau(A, C) = 15 - 3 = 12. \tag{10}$$

With this notation, $\tau(X, Y) > 0$ means that candidate X received the larger vote, so X beats Y in the paired comparison. As such, options X, Y, Z define the transitive relationship $X \succ Y \succ Z$ if and only if X beats Y, Y beats Z, and X beats Z, which is

$$\tau(X, Y) > 0, \ \ \tau(Y, Z) > 0, \ \ \textit{and} \ \ \tau(X, Z) > 0. \tag{11}$$

While the Eq. 9 example satisfies the Eq. 11 inequalities ensuring a transitive outcome, the Eq. 10 values go a powerful step further by satisfying the *equality*

$$\tau(A, B) + \tau(B, C) = \tau(A, C).$$

Extending far beyond usual transitivity conditions (Eq. 11), these particular tallies resemble measurements along a line where the (signed) distance from A to B plus the (signed) distance of B to C equals the (signed) distance from A to C.[1] Demonstrating my unimaginative choice of words, I call this stronger form of transitivity *"strongly transitive*."

Definition 1 (Saari 2000, 2008) The pairwise rankings over alternatives $\{X_j\}_{j=1}^{N}$ satisfy "strong transitivity" ($\mathcal{ST}$) if all triplets satisfy the equality

[1] The "signed" qualifier means that the direction in which the measurements are taken is being considered.

$$\tau(X_i, X_j) + \tau(X_j, X_k) = \tau(X_i, X_k). \tag{12}$$

The above $N = 3$ illustration suggests using the $\mathcal{RWC}$ balancing condition to determine whether or not a profile is $\mathcal{ST}$. Theorem 1 proves that this is the case.

Theorem 1. A N-candidate profile $\mathbf{p}$ satisfies $\mathcal{ST}$ if and only if for each $\mathcal{RWC}$, the number of voters with rankings in this $\mathcal{RWC}$ agrees with the number of voters with preference rankings in the $\mathcal{RWC}$ defined by reversing the original initial ranking.

Stated in words, a $\mathcal{ST}$ profile is "balanced" with respect to all $\mathcal{RWC}$s. It is easy to verify whether a given profile satisfies this condition. To illustrate, consider

Number	Ranking	Number	Ranking
3	$A \succ B \succ C \succ D \succ E$	2	$C \succ A \succ B \succ E \succ D$
2	$C \succ D \succ E \succ A \succ B$	2	$E \succ B \succ A \succ C \succ D$
5	$C \succ B \succ A \succ E \succ D$		

(13)

To check whether this profile is $\mathcal{ST}$, start with the $\mathcal{RWC}$ defined by the first ranking $A \succ B \succ C \succ D \succ E$ and the $\mathcal{RWC}$ defined by its reversal $E \succ D \succ C \succ B \succ A$. Notice that the first and second rankings listed on the left belong to the first $\mathcal{RWC}$, while the last one on the left belongs to the reversed $\mathcal{RWC}$. Each configuration has five voters using its rankings, so the balanced condition is satisfied with this particular $\mathcal{RWC}$ and its reversal.

To examine the remaining rankings (listed on the right), consider the $\mathcal{RWC}$s defined by the first ranking and its reversal. Each ranking belongs to a different one of these $\mathcal{RWC}$s. As the number of voters with rankings from each $\mathcal{RWC}$ agree, this profile is $\mathcal{ST}$. Checking Eq. 12 with, say, triplet $\{A, C, E\}$, a count shows that $\tau(A, C) = 5 - 9 = -4$, $\tau(C, E) = 12 - 2 = 10$, $\tau(A, E) = 10 - 4 = 6$, so, as required, $\tau(A, C) + \tau(C, E) = \tau(A, E)$.

Theorem 1 makes it easy to create $\mathcal{ST}$ profiles; use a balanced number of voters with preferences coming from selective $\mathcal{RWC}$s. As an $N = 4$ example, start with 5 voters preferring $A \succ B \succ C \succ D$, and 3 preferring $C \succ D \succ A \succ B$ (both come from the $\mathcal{RWC}$ defined by $A \succ B \succ C \succ D$) and balance this with 7 preferring $C \succ B \succ A \succ D$ and one prefers $A \succ D \succ C \succ B$ (coming from the $\mathcal{RWC}$ defined by the reversed $D \succ C \succ B \succ A$). Adding another ranking with 3 preferring $C \succ A \succ D \succ B$ requires a balance with 3 voters having rankings in the $\mathcal{RWC}$ defined by $B \succ D \succ A \succ C$, so let 2 have $D \succ A \succ C \succ B$ and 1 have $A \succ C \succ B \succ D$. By construction this 22 voter profile satisfies the balancing condition with respect to all $\mathcal{RWC}$s, so it is $\mathcal{ST}$. To illustrate Eq. 12 with, say, $\{D, A, B\}$, a count shows that $\tau(D, A) = 5 - 17 = -12$, $\tau(A, B) = 15 - 7 = 8$, $\tau(D, B) = 9 - 13 = -4$, so $\tau(D, A) + \tau(A, B) = \tau(D, B)$ as required.

3.2 Dividing Profiles into "Nicely" and "Badly" Behaved Parts

Although simple, these constructions have surprisingly strong consequences: They introduce a tool to completely characterize everything that can happen with paired comparisons.

The tool is created is by dividing the domain (the $N!$-dimensional profile space) into two orthogonal parts: the $\mathcal{ST}$ and the $\mathcal{RWC}$ profiles. This means that all profiles consist of a $\mathcal{ST}$ and a $\mathcal{RWC}$ component; *there is nothing else*. Because, as developed below, the $\mathcal{ST}$ profiles never cause problems, it follows that the $\mathcal{RWC}$ portion of profiles are completely responsible for the several centuries of mysteries, debate, and problems. Stated informally, Theorem 2 divides profiles into "nicely behaved" ($\mathcal{ST}$) and "badly behaved" ($\mathcal{RWC}$) parts.

Theorem 2. For any $N \geq 3$, the span of the ranking wheel configurations and the strongly transitive configurations covers the full profile space. The space of $\mathcal{RWC}$ profiles has dimension $\frac{(N-1)!}{2}$, while that of $\mathcal{ST}$ profiles has dimension $(2N-1)\frac{(N-1)!}{2}$.

Starting in the next subsection, consequences of Theorem 2 are explored. The first topic is to examine Nakamura's number.

3.2.1 Nakamura's Number

As nothing goes wrong with $\mathcal{ST}$ portions of a profile, it follows from Theorem 2 that all possible problems with paired comparisons must be caused by the $\mathcal{RWC}$ portions. To appreciate how to use this information, return to the earlier problem where the winner of a paired comparison requires 72 of the 84 possible votes. The goal is to determine the minimum number of alternatives that are needed to generate a cycle; it follows from Theorem 2 that the answer must involve $\mathcal{RWC}$s.

According to Theorem 2, any example causing such a cycle must involve a $\mathcal{RWC}$. So start with a $N = 10$ $\mathcal{RWC}$; here the paired comparisons have tallies of $N-1$:1, or 9:1. As this 9:1 tally does not suffice, several copies of this $\mathcal{RWC}$ are needed; the key step is to determine how many. This is easy; to reach a tally of 72, we need $\frac{72}{9} = 8$ copies. As each $\mathcal{RWC}$ involves ten voters, these eight copies of the $\mathcal{RWC}$ define the preferences for $8 \times 10 = 80$ voters. The preferences of these 80 voters assure the cycle, so the last four voters can be assigned any preference ranking. Thus, with ten alternatives, a cycle can be created.

Now try the smaller number of seven alternatives. Here the $\mathcal{RWC}$ cycle has 6:1 tallies, which, again, requires using several copies of this $\mathcal{RWC}$. The necessary number is the smallest integer equal to or greater than $\frac{72}{6} = 12$, so twelve copies are needed to generate a cycle. These copies assign the preferences for $7 \times 12 = 84$ voters. As 84 is the precise number of voters that is available, a cycle can be constructed.

Next try six alternatives; as the ranking wheel configuration defines a cycle with 5:1 tallies, the first integer equal to or greater than $\frac{72}{5}$ is 15, so 15 copies of this $\mathcal{RWC}$ are needed to create a cycle. These 15 copies assign preferences to $6 \times 15 = 90$ voters. But as there are only 84 voters, this cannot be done. Thus, with six or fewer alternatives, a $q = 72$ voter cycle cannot be created.

In general, with a q-rule with n voters, a N-alternative ranking wheel yields the tallies of $N - 1$:1. Thus $\frac{q}{N-1}$ copies of this $\mathcal{RWC}$ are required to create an example. These copies define the preferences for $N\frac{q}{N-1}$ voters. To ensure that a cycle can be created, at least this number of voters are needed; i.e., $N\frac{q}{N-1} \leq n$, or $N \geq \frac{n}{n-q}$. Stated in another manner, if $N < \frac{n}{n-q}$, a q-rule cycle cannot be created. This "first integer greater than or equal to $\frac{n}{n-q}$" threshold value is the Nakamura number. For the example, the threshold value is $\frac{84}{12} = 7$.

Completing a proof of this result is simple; according to Theorem 2, any issue concerning cycles involves ranking wheel configurations, and only ranking wheel configurations. No other kind of profile can possibly create or enhance tallies in a cycle. The only remaining concern is show that using copies of a single $\mathcal{RWC}$, rather than combinations, maximizes the paired comparison tallies in a cycle; this is a simple computation.

Notice how this analysis introduces a new interpretation of Nakamura's number. Namely, the Nakamura number merely specifies the number of copies of a $\mathcal{RWC}$ that are required to create a profile that will have a cycle for a specified q-rule. This number of copies determines the needed number of voters, which limits the number of admissible alternatives.

3.2.2 Dividing a Profile

When using Theorem 2, it is highly unlikely that a given profile is strictly $\mathcal{ST}$ or $\mathcal{RWC}$. A profile is like a person; not all bad nor all good, but some combination. The Eq. 2 profile (illustrating that the Borda winner need not be the Condorcet winner) is neither $\mathcal{ST}$, nor $\mathcal{RWC}$; it is a combination. Indeed, adding the profiles $\mathbf{P}_1$ and $\mathbf{P}_2$ from the following Eq. 14 equals the Eq. 2 profile. As it is easy to check, $\mathbf{P}_1$ is $\mathcal{ST}$ while $\mathbf{P}_2$ is a $\mathcal{RWC}$, so this example illustrates the Theorem 2 assertion.

$$\mathbf{P}_1 = \begin{array}{cc|cc} \textbf{No.} & \textbf{Rank} & \textbf{No.} & \textbf{Rank} \\ \hline 6 & A \succ B \succ C & 0 & C \succ B \succ A \\ 4 & A \succ C \succ B & 2 & B \succ C \succ A \\ 1 & C \succ A \succ B & 5 & B \succ A \succ C \end{array}, \quad \mathbf{P}_2 = \begin{array}{cc} \textbf{No.} & \textbf{Rank} \\ \hline 5 & A \succ C \succ B \\ 5 & C \succ B \succ A \\ 5 & B \succ A \succ C \end{array} \tag{14}$$

Staying with Eq. 14 notice how the paired comparison tallies of the $\mathcal{ST}$ portion $\mathbf{P}_1$ are $A \succ B$, $A \succ C$ with respective tallies of 11:7 and 15:3. Thus A, not B (as in

Eq. 2), is the Condorcet winner for the profile's $\mathbf{P}_1$ portion. (The remaining paired outcome is $B \succ C$ with tally 13:5.) Over this $\mathcal{ST}$ portion, the Borda ranking is $A \succ B \succ C$ with tally 26:20:8. Thus, with the $\mathbf{P}_1$ portion of the profile, both methods completely agree.

To understand why Borda and paired comparisons can differ, notice that the paired comparisons coming from the $\mathcal{RWC}$ portion $\mathbf{P}_2$ define the cycle $B \succ A, A \succ C, C \succ B$, each with a 10:5 tally, while the Borda ranking for $\mathbf{P}_2$ is the complete tie $A \sim B \sim C$ with common tally of 15. For this example, then, the Sect. 2.1 difference in voting outcomes is *strictly caused by the* $\mathcal{RWC}$ *portion of the profile*. More specifically, the distorted 10:5 tally from the cyclic $\mathcal{RWC}$ portion is what changes the $\{A, B\}$ outcome to $B \succ A$ with the earlier reported tally $(7 + 10)$:$(11 + 5)$. In other words, the force of these $\mathcal{RWC}$ terms is what replaces A with B as the Condorcet winner. In contrast, Borda is immune to this $\mathcal{RWC}$ portion of a profile (where no candidate is "better than" any other); that is, $\mathcal{RWC}$ components never effect the Borda ranking.

To summarize, the cyclic effect of the $\mathcal{RWC}$ portion of the Eq. 2 profile is what causes B to become the Condorcet winner and create the conflict with the Borda ranking. As reported in Theorem 3, this always happens; *all differences* between the Borda and paired comparison rankings and tallies are strictly due to $\mathcal{RWC}$ portions of profiles. In turn, this means that any comparison between, say, Borda and Condorcet reduces to emphasizing the value of being affected by, or not being affected by $\mathcal{RWC}$ terms.

3.2.3 Some Vector Analysis

Not all profiles split as nicely as Eq. 2 into $\mathcal{ST}$ and $\mathcal{RWC}$ components (Eq. 14). As a surprising example, the unanimity profile with a single voter's preference $A \succ B \succ C$ has values $\tau(A, B) = \tau(B, C) = \tau(A, C) = 1$, so Eq. 12 cannot be satisfied; i.e., the unanimity profile is *not* $\mathcal{ST}$.[2] Moreover, this profile consists of a single voter, so it cannot be represented (as in Eq. 14) with integer numbers of voters with $\mathcal{RWC}$ and $\mathcal{ST}$ preferences. Instead, "fractional numbers" of voters with different characteristics must be used. Actually, such a use of fractions is standard; e.g., an academic could have a $\frac{2}{3}$ appointment in mathematics and a $\frac{1}{3}$ appointment in economics. A similar "accounting" construction is needed for Theorem 2.

This description requires converting profiles from the traditional listing of preferences into a vector representation. On a first reading of this article, a reader can safely skip this mathematical subsection and jump to Sect. 3.3. But the reader

[2] The reader may be surprised and bothered by this comment; I was when I discovered this result. But, as described near the end of this section, the fact the unanimity profile is not $\mathcal{ST}$ explains some mysteries. As an illustration, if all that is known about an unanimity profile is that $\tau(C, B) = \tau(C, A) = 1$, then C is top-ranked, but who is second ranked? As shown below, this information is provided by the $\mathcal{ST}$ portion of this profile; this $\mathcal{ST}$ information is lost with the actual profile because of the $\mathcal{RWC}$ influence.

is strongly encouraged to return to review these tools because they are needed to fully understand and use Theorem 2.

With $N!$ ways to rank N alternatives, the profile space resides in $\mathbb{R}^{N!}$, so profiles can be represented as vectors in $\mathbb{R}^{N!}$. To do so, identify each $\mathbb{R}^{N!}$ axis with a ranking. In this manner, a $\mathbb{R}^{N!}$ vector lists how many voters have each preference ranking. Illustrating with $N = 3$, list the $3! = 6$ rankings in the order

$$A \succ B \succ C,\, A \succ C \succ B,\, C \succ A \succ B,\, C \succ B \succ A,\, B \succ C \succ A,\, B \succ A \succ C,$$

where an adjacent pair is reversed to move from one ranking to the next one. (For a "geometric" explanation of this ordering choice, see Saari 2010.) In this manner, the Eq. 9 profile has the representation $(6, 4, 1, 0, 2, 5)$, the $\mathcal{RWC}$ triplet Eq. 6 becomes $\mathbf{C}_{ABC} = (1, 0, 1, 0, 1, 0)$ and Eq. 7 is $\mathbf{C}_{ACB} = (0, 1, 0, 1, 0, 1)$.

These vectors play a critical role when determining the Theorem 2 dimensions. As an illustration, $\mathbf{C}_{ABC} + \mathbf{C}_{ACB} = \mathbf{E}_6 = (1, 1, 1, 1, 1, 1)$. As profile $\mathbf{E}_6$ has the same number of voters of each type, all paired comparisons end in ties; i.e., profile $\mathbf{E}_6$ is $\mathcal{ST}$. This creates a slight mathematical problem because to place $\mathcal{ST}$ and $\mathcal{RWC}$ components in orthogonal subspaces, such a sum (of two $\mathcal{RWC}$ terms equalling a $\mathcal{ST}$ term) must be prohibited. A way to avoid this difficulty is to use $\mathbf{C} = \mathbf{C}_{ABC} - \mathbf{C}_{ACB} = (1, -1, 1, -1, 1, -1)$ to represent this $\mathcal{RWC}$; vector $\mathbf{C}$ is orthogonal to $\mathbf{E}_6$. (See Saari (2000) or (2008), Chap. 4 for more arguments and an explanation of how to interpret "profiles" with negative components.)

To illustrate $\mathbf{C}$, recall that a three-alternative profile is $\mathcal{ST}$ if it satisfies the balanced condition; this requirement is equivalent to requiring the scalar product of the profile's vector representation with $\mathbf{C}$ to equal zero. Illustrating with the Eq. 9 example of $(6, 4, 1, 0, 2, 5)$, its scalar product with $\mathbf{C}$ is

$$\begin{aligned}((6, 4, 1, 0, 2, 5),\quad & (1, -1, 1, -1, 1, -1)) = \\ & [(6)(1) + (4)(-1) + (1)(1) + (0)(-1) + (2)(1) + (5)(-1)] = 0.\end{aligned}$$

But with profile $\mathbf{p} = (3, 1, 0, 1, 2, 0)$, the inner product $(\mathbf{p}, \mathbf{C}) = 5 - 2 = 3$ proves that this $\mathbf{p}$ is *not* $\mathcal{ST}$. The goal is to find $\mathbf{p}$'s decomposition. The first step is to determine how much of $\mathbf{p}$ is in the direction of the $\mathcal{RWC}$ vector $\mathbf{C}$. To do so, recall from vector analysis that a "direction" has length one. Thus this three-alternative $\mathcal{RWC}$ "direction" is $\tilde{\mathbf{C}} = \frac{1}{\sqrt{6}}\mathbf{C} = \frac{1}{\sqrt{6}}(1, -1, 1, -1, 1, -1)$. Also recall that the amount of a vector in a given direction is determined by the scalar product; here it is

$$(\mathbf{p}, \tilde{\mathbf{C}}) = \frac{1}{\sqrt{6}}(\mathbf{p}, \mathbf{C}) = \frac{1}{\sqrt{6}}3,$$

where the $(\mathbf{p}, \mathbf{C}) = 3$ computation comes from the first line of this paragraph. This means that the amount of profile $\mathbf{p}$ in the $\mathcal{RWC}$ direction is given by

$$\frac{3}{\sqrt{6}}\tilde{\mathbf{C}} = \frac{3}{\sqrt{6}}\left[\frac{1}{\sqrt{6}}\mathbf{C}\right] = \frac{1}{2}\mathbf{C} = \left(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}\right). \tag{15}$$

The remaining part of $\mathbf{p}$, which defines the $\mathcal{ST}$ portion, is given by

$$\mathbf{p} - \frac{1}{2}\mathbf{C} = (3, 1, 0, 1, 2, 0) - \left(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}\right) = \left(\frac{5}{2}, \frac{3}{2}, -\frac{1}{2}, \frac{3}{2}, \frac{3}{2}, \frac{1}{2}\right). \tag{16}$$

Thus, in analyzing profile $\mathbf{p}$, the Eq. 15 portion identifies what causes problems while the Eq. 16 part captures the well-behaved portion.

Similarly, the earlier unanimity profile $\mathbf{u} = (1, 0, 0, 0, 0, 0)$ has the $\mathcal{RWC}$ portion $\frac{1}{6}\mathbf{C}$. Thus $\mathbf{u} - \frac{1}{6}\mathbf{C} = (\frac{5}{6}, \frac{1}{6}, -\frac{1}{6}, \frac{1}{6}, -\frac{1}{6}, \frac{1}{6})$ is the $\mathcal{ST}$ portion with τ tallies for $A \succ B, B \succ C, A \succ C$ of, respectively, $\tau(A,B) = \frac{2}{3}, \tau(B,C) = \frac{2}{3}, \tau(A,C) = \frac{4}{3}$. Notice how this $\mathcal{ST}$ portion with $\tau(A,C) = \frac{4}{3}$ accurately captures the fact that A is more strongly preferred over C then A is over B. This difference is missing from the earlier misleading $\tau(A,B) = \tau(B,C) = \tau(A,C) = 1$, which is a consequence of the unanimity profile's $\mathcal{RWC}$ component.

For $N > 3$, each $\mathcal{RWC}$ defines a vector version of the above $\mathbf{C}$. The way this is done is to list the rankings of the $\mathcal{RWC}$ in a vector form. (This choice depends on which coordinates of this $N!$ vector represent which rankings.) Next, find the $\mathcal{RWC}$ defined by reversing one of the rankings of the original choice; as with $\mathbf{C}$, represent each of these rankings in the vector with -1's. In this manner, the dimension of the space spanned by ranking wheel configurations is cut in half to $\frac{(N-1)!}{2}$. Thus, by being orthogonal to the $\mathcal{RWC}$, the space of $\mathcal{ST}$ profiles has dimension $N! - \frac{(N-1)!}{2} = (2N-1)\frac{(N-1)!}{2}$, which proves Theorem 2. (But most $\mathcal{ST}$ profiles define a complete tie for pairs; see Theorem 6.)

3.3 Consequences of Strong Transitivity

Some of the advantages gained from the Theorem 2 decomposition are indicated by the following theorems. The first one explains the problem raised in Sect. 2.1 where Borda and Condorcet winners can differ. As suggested when analyzing the Eq. 2 profile in terms of Eq. 14, Theorem 3 asserts that all possible differences between these rules are due to the $\mathcal{RWC}$ portion of a profile.

Theorem 3. (Saari 2000) For any $N \geq 3$ and with $\mathcal{ST}$ profiles, the election rankings for the Borda Count and majority votes over pairs always agree. Even stronger, the tallies agree in the following strong sense: Let *Borda*(X) be X's Borda tally. For each N and any two alternatives X and Y, it always is true that

$$Borda(X) - Borda(Y) = \frac{N}{2}\tau(X, Y). \tag{17}$$

If a profile is not strictly $\mathcal{ST}$, Eq. 17 is not satisfied for all pairs of alternatives; some Borda and paired comparison tallies must differ. All possible differences between Borda and paired comparison rankings and tallies are due to a profile's $\mathcal{RWC}$ components.

3.3.1 Borda Versus Condorcet

Equation 17 is surprising; it requires the Borda and paired comparison tallies for the $\mathcal{ST}$ portion of a profile to always agree. To illustrate the assertion, recall that paired comparison tallies from $\mathbf{P}_1$ (Eq. 14) have $A \succ B$, $A \succ C$, and $B \succ C$ with the respective tallies of 11:7, 15:3, and 13:5. Thus the tally differences are

$$\tau(A, B) = 11 - 7 = 4, \quad \tau(A, C) = 15 - 3 = 12, \quad \tau(B, C) = 13 - 5 = 8.$$

The Borda ranking for $\mathbf{P}_1$ is $A \succ B \succ C$ with tally 26: 20: 8. As required by Eq. 17

$$\begin{aligned} Borda(A) - Borda(B) &= 26 - 20 = 6 = \tfrac{3}{2}\tau(A, B) \\ Borda(A) - Borda(C) &= 26 - 8 = 18 = \tfrac{3}{2}\tau(A, C), \\ Borda(B) - Borda(C) &= 20 - 8 = 12 = \tfrac{3}{2}\tau(B, C). \end{aligned}$$

An $N = 5$ illustration comes from the Eq. 13 profile where the Borda ranking is $C \succ A \sim B \succ E \succ D$ with the 44:34:34:19:9 tally. For this example, the values

$$\tau(A, C) = 5 - 9 = -4, \quad \tau(C, E) = 12 - 2 = 10, \quad \tau(A, E) = 10 - 4 = 6$$

are computed above. The equalities required by Eq. 17 are

$$\begin{aligned} Borda(A) - Borda(C) &= 34 - 44 = -10 = \tfrac{5}{2}\tau(A, C) \\ Borda(C) - Borda(E) &= 44 - 19 = 25 = \tfrac{5}{2}\tau(C, E), \\ Borda(A) - Borda(E) &= 34 - 19 = 15 = \tfrac{5}{2}\tau(A, E). \end{aligned}$$

To illustrate how Theorem 3 provides new insights into troubling issues, return to the main criticism of the Borda Count that it need not elect the Condorcet winner. According to Theorem 3, when the Borda and Condorcet winners differ, the reason is completely due to a profile's $\mathcal{RWC}$ portion; Borda ignores these terms while Condorcet is influenced by them. This suggests that the standard criticism is stated in the wrong order; rather than using a criterion (Condorcet's approach) that fails to appropriately handle $\mathcal{RWC}$ components, perhaps a more accurate criticism is that

A problem with majority votes over pairs is that they need not elect the Borda winner. A problem with the Condorcet winner is that it need not be the Borda winner.

To more fully determine which criticism is correct, it is necessary to understand the effect a profile's $\mathcal{RWC}$ portion has on the outcome. This is the theme of Sect. 3.4.

3.3.2 Analyzing Other Paired Comparison Results

Theorem 3 indicates how to analyze any concern involving paired comparisons as determined by profiles. (This requirement of using information only from "profiles" is removed in Sect. 4.) Namely, prove that the $\mathcal{ST}$ portion of a profile avoids all difficulties, while the $\mathcal{RWC}$ portion causes all of the complexities.

This approach can be used to analyze any paired comparison method, such as Kemeny's rule (Kemeny 1959). The ones selected here are widely known: Arrow's Impossibility Theorem (Arrow 1951) and Sen's "Impossibility of a Paretian liberal" Saari (1970a, b). What ensures that these two results are subject to Theoem 2 is that both assertions require the group decisions to be constructed in terms of paired comparisons.

In Arrow's result, the two conditions imposed on the rule are that

(1) (Pareto) If everyone ranks a *pair* in the same manner, this common ranking is the pair's societal ranking.
(2) (IIA) A *pair's* societal ranking depends only on how the voters rank this particular pair; all other information is irrelevant.

Sen keeps Arrow's Pareto condition; instead of the general IIA requirement, he uses a special IIA case where at least two agents can determine the societal ranking of specified pairs of options. Both results lead to the same negative conclusion that no such rule exits. (Reader not familiar with Arrow's and Sen's theorems can find information in many references where, for obvious reasons, my preference is Saari 2008, Chap. 2.)

By requiring paired comparisons, both theorems can be analyzed with Theorem 2. Following the above template of how to use this decomposition, the main message of the following Theorem 4 is that the negativity of these seminal assertions is strictly caused by the $\mathcal{RWC}$ portion of a profile. An immediate corollary, then, is that by removing this portion from a profile, all difficulties registered by these two results disappear.

Theorem 4. For any $N \geq 3$, if the profiles are $\mathcal{ST}$, then the Borda Count satisfies Arrow's (Arrow 1951) conditions for a social choice function. Similarly, if only $\mathcal{ST}$ profiles are admitted, then Sen's conditions stated for his Paretian Liberal Theorem (Sen 1970a) never admit cycles.

Stating this theorem in different terms, the negativity of the Arrow and Sen theorems are direct consequences of using paired comparisons. This choice forces the decision rule to be influenced by $\mathcal{RWC}$ portions of profile, which then introduce cyclic effects. But by removing these disruptive components, it follows from Theorem 4 that both assertions lose their mystery; the associated conclusions are positive and readily acceptable.

A closely related goal is to find ways to replace Arrow's and Sen's negative assertions with positive conclusions. To do so, first notice what causes their negative assertions: To avoid difficulties, a rule must be able to recognize and correct for $\mathcal{RWC}$ portions of a profile. Because these $\mathcal{RWC}$ portions of a profile are *not* local in nature, the rule *cannot* be restricted to using information just about individual pairs; it must be able to recognize more global connections among the pairs. But Arrow and Sen explicitly prohibit a rule from doing so. As such, to replace their assertions with positive statements requires finding ways to allow a rule to identify and negate the impact of the $\mathcal{RWC}$ portion of a profile. (This is done in Saari 2008, Chap. 2.) But first, a deeper understanding of the negative effects of $\mathcal{RWC}$'s must be developed; this is the theme of Sect. 3.4.

3.4 Information Lost by Ranking Wheel Configurations

As we now know, complexities experienced by paired comparisons are strictly due to $\mathcal{RWC}$ portions of profiles. These difficulties reflect the fact that crucial but available information about the global structure is being ignored. A goal is to identify some of the costs of ignoring this information.

A hint of what goes wrong comes from the discussion at the beginning of Sect. 3 with the introduction of the ranking wheel; paired comparisons ignore the global structure of the $\mathcal{RWC}$ suggesting that the outcome should be a tied vote to create a cyclic outcome. A surprise is that the actual harm caused by this $\mathcal{RWC}$ portion of a profile is far more severe than normally appreciated; to describe what it is, an argument developed in Saari (2008, Chap. 2) is slightly modified.

3.4.1 International Cooperation

Suppose an international committee is being formed where one country from each of the three groups is to be selected.

Group 1	Group 2	Group 3
Argentina	Brazil	Chile
Denmark	Estonia	Finland

(18)

Suppose the majority vote selection is made by three officials; each official chooses one country from each of the three groups. Would the outcome of {Denmark, Estonia, Finland}, where each vote is 2:1, reflect the intent of the officials?

To analyze this question, examine all possible supporting profiles; ignoring names of the voters, there are five possible profiles. Using only a country's first initial, the five are

Profile	Voter 1	Voter 2	Voter 3
1	A, B, C	D, E, F	D, E, F
2	D, B, C	A, E, F	D, E, F
3	A, E, C	D, B, F	D, E, F
4	A, B, F	D, E, C	D, E, F
5	A, E, F	D, B, F	D, E, C

(19)

Profiles 1 to 4 share the feature where the preferences of the first two voters create a tie, which is then broken by the third voter. The {Denmark, Estonia, Finland} outcome is reasonable in these first four of the five cases, so it follows that the conclusion is appropriate for at least 80 percent of the profiles. About the only way to justify the last possibility (profile 5) is to note that each country received two of the three votes.

Suppose that the outlier profile 5 is the actual one. If the officials shared an "intercontinental participation" goal of including representatives from both Northern Europe and South America (which is consistent with how each official voted), everyone would be disappointed with the outcome. On the other hand, criticizing the rule appears to be unfair; after all, it is impossible to anticipate what kinds of side conditions might be intended. Consequently, if this "intercontinental participation" constraint is truly desired, it must be built into the voting rule. If it is not, there is no reason to expect the rule to recognize or honor this side condition.

3.4.2 Loss of Transitivity

My main point is to illustrate that should a constraint (e.g., intercontinental participation) be desired, it must be built into a voting rule. If it is not, there is no reason to expect the rule to recognize that the voters want to satisfy this condition, nor to expect the rule's outcomes to satisfy it.

Eq. 18 involves pairwise voting, so Theorem 2 can be used to examine this issue. To assist the analysis, directly translate this example into a voting setting over alternatives α, β, γ by identifying D with $\alpha \succ \beta$ so A is identified with $\beta \succ \alpha$, E with $\beta \succ \gamma$ so B is identified with $\gamma \succ \beta$, and F with $\gamma \succ \alpha$ so C is identified with $\alpha \succ \gamma$. This "name-change" converts the {Denmark, Estonia, Finland} conclusion into the $\alpha \succ \beta$, $\beta \succ \gamma$, $\gamma \succ \alpha$ cycle where each tally is 2:1. The fact the outcome is identified with a cycle suggests an unexpected, hidden involvement of a $\mathcal{RWC}$ component.

Just a "name-change" converts Eq. 19 into the equivalent listing:

Profile	**Voter 1**	**Voter 2**	**Voter 3**
1	$\beta \succ \alpha, \gamma \succ \beta, \alpha \succ \gamma$	$\alpha \succ \beta, \beta \succ \gamma, \gamma \succ \alpha$	$\alpha \succ \beta, \beta \succ \gamma, \gamma \succ \alpha$
2	$\alpha \succ \gamma \succ \beta$	$\beta \succ \gamma \succ \alpha$	$\alpha \succ \beta, \beta \succ \gamma, \gamma \succ \alpha$
3	$\beta \succ \alpha \succ \gamma$	$\gamma \succ \alpha \succ \beta$	$\alpha \succ \beta, \beta \succ \gamma, \gamma \succ \alpha$
4	$\gamma \succ \beta \succ \alpha$	$\alpha \succ \beta \succ \gamma$	$\alpha \succ \beta, \beta \succ \gamma, \gamma \succ \alpha$
5	$\beta \succ \gamma \succ \alpha$	$\gamma \succ \alpha \succ \beta$	$\alpha \succ \beta \succ \gamma$

(20)

As only names have been changed, features that are true for Eq. 19 translate into features that are true for Eq. 20. By comparing the lists, we find that:

(1) A listing satisfies an "intercontinental" condition if and only if its translation is a "transitive ranking." Any listing where all candidates come from the same continent is equivalent to a cyclic ranking of the alternatives.
(2) The fifth profile from Eq. 20 is a $\mathcal{RWC}$; thus the problems experienced by the actual Eq. 18 example profile reflect properties of $\mathcal{RWC}$s.
(3) What allows the Eq. 18 outcome to be viewed as reasonable is that it is appropriate for 80 "appropriate" because it is supported by 80
(4) It is unreasonable to expect the paired comparisons used in Eq. 18 to reflect the "intercontinental constraint" because the condition is not part of the voting rule. Similarly, it is unreasonable to expect paired comparisons to yield transitive outcomes for a $\mathcal{RWC}$ because this "transitivity constraint" is not built into the voting rule.

In other words, *the combination of paired comparisons and $\mathcal{RWC}$ negates the crucial assumption that voters have transitive preferences!* By assumption, the voters do have transitive preferences; problems are caused because paired comparison rules cannot identify nor use this fact. Instead, a paired comparison rule interprets a profile with a sizable $\mathcal{RWC}$ component as reflecting the views of non-existent voters with cyclic preferences. The information lost through $\mathcal{RWC}$ portions of profiles, then, is the individual rationality of the voters; $\mathcal{RWC}$ components evaluated by paired comparisons trash the crucial assumption that voters have transitive preferences.

Paired comparisons, then, sever all connections among parts, whether intended or not, the parts are treated as separate, independent entities without connections. In a very real sense, the rule finds an appropriate answer for most profiles that consist of these particular parts; this is reflected by the outcomes for Eqs. 19 and 20 reflecting 80 possible associated profiles. In general, to find the set of all profiles, ignore connections and merely assemble the parts in all possible ways. Let me recommend that the reader use this argument to analyze the Assistant Professor example of Eq. 1. Everything translates immediately, and the analysis is particularly telling. In particular, an emphasis on the parts means that the approach cannot distinguish the actual Eq. 1 setting from others where the data is clear, inconsistent, or nonsensical (e.g., a judge votes "No" on each part, and "Yes" on tenure).

3.4.3 Returning to Arrow's Theorem and Other $\mathcal{RWC}$ Consequences

The above discussion explains Arrow's assertion. By requiring societal decisions to be determined with paired comparisons, his result is held captive to $\mathcal{RWC}$ portions of profiles, which in turn dismiss the crucial assumption that voters have transitive preferences. Remember, voters *do have transitive preferences.* But when paired comparisons encounter $\mathcal{RWC}$ components, the rule treats them as reflecting the wishes of nonexistent voters who have cyclic preferences.

This comment immediately suggests how to circumvent the negativity of Arrow's result. Recall, if the "intercontinental participation" requirement is desired, it must be built into the decision rule. Similarly, if transitive outcomes are desired, this requirement must be built into the decision rule. Namely, this individual rationality intent must be built into Arrow's assumptions. There are many ways to do so; my favorite is to recognize that with a transitive ranking, it is possible to specify the number of alternatives that separate any two in a ranking. But with cyclic preferences, this is impossible. So, my choice is to modify Arrow's IIA condition (Sect. 3.3.2) with the following:

Definition 2. (Saari 1995, 2008) The "Intensity of Independence of Irrelevant Alternatives" (IIIA) is where the societal outcome for each pair is determined by how each voter ranks this pair *along with the number of alternatives that separates this pair in the voter's transitive ranking.*

To illustrate with the $A \succ B \succ C \succ D$ ranking, when finding the societal ranking for $\{A, D\}$, Arrow's IIA condition permits using only the $[A \succ D]$ information. My IIIA condition allows the rule to use $[A \succ D, 2]$ to reflect that two alternatives separate A and D in this ranking. The IIIA property, then, counters the negative effects of the $\mathcal{RWC}$ components that force Arrow's negative assertion. This can be seen with the following:

Theorem 5. (Saari 2008) With $N \geq 3$ alternatives and at least two voters, assume that the voters have complete, transitive rankings of the alternatives and that the societal outcome is to be a complete, transitive ranking. A rule that does so while satisfying Pareto and IIIA is the Borda Count.

The Borda Count satisfies Theorem 5 by introducing a new way to compute pairwise votes in an $\{X, Y\}$ election. Namely, when tallying a ballot with the $[X \succ Y, k]$ information, assign $k+1$ points to X and zero to Y. In this assignment, X receives 1 point by being ranked above Y on the ballot and k bonus points by being preferred significantly more strongly than Y.

This tallying approach eliminates $\mathcal{RWC}$ problems. The Eq. 6 profile, for instance, reflects $\mathcal{RWC}$ difficulties with its majority vote cycle. But with this new tallying method, the $\{A,B\}$ information is $[A \succ B, 0]$, $[B \succ A, 1]$, and $[A \succ B, 0]$. Thus, A's tally is $1 \times 1 + 1 \times 1 = 2$, while B's tally is $1 \times (1+1) = 2$ to create a tie. Similarly, the rankings for all pairs end in ties, so the cycle is averted. The fact this rule eliminates the effects of $\mathcal{RWC}$ is what makes positive assertions possible. (I leave it as an exercise to show that the majority vote rankings for a $\mathcal{ST}$ profile agree with the rankings determined by this new approach.)

To further illustrate with the Eq. 2 profile and the $\{A,B\}$ outcome (where the Borda and usual majority vote rankings disagree), the column on the left has 6 voters with $[A \succ B, 0]$, 9 with $[A \succ B, 1]$, and 1 with $[A \succ B, 0]$, while the column on the right has 5 voters with $[B \succ A, 0]$, 2 with $[B \succ A, 1]$, and 10 with $[B \succ A, 0]$. With this new paired comparison rule, then, A receives $6 \times 1 + 9 \times 2 + 1 \times 1 = 25$ while B receives $5 \times 1 + 2 \times 2 + 10 \times 1 = 19$ leading to the $A \succ B$ outcome that is consistent with the $A \succ B \succ C$ Borda ranking. Even more, the tally difference between A and B is $25 - 19 = 6$ which agrees with $Borda(A) - Borda(B) = 41 - 35 = 6$. This agreement is no accident; it always occurs (which is not difficult to prove): this is why Borda satisfies Theorem 5.[3]

In a similar way, Sen's Theorem also can be explained in terms of the $\mathcal{RWC}$s. His result normally is interpreted as identifying a conflict between the Pareto condition that involves all agents and allowing specific agents to make personal decisions over specified pairs. But the real source of Sen's negative conclusion is that $\mathcal{RWC}$ portions of a profile negate the crucial assumption that voters have transitive preferences. It is interesting how, when examining the effect of $\mathcal{RWC}$s on Sen's result, radically different conclusions emerge. One is that rather than modeling normal types of interactions, Sen's result actually captures *dysfunctional settings* where the personal actions of some agents impose strongly negative burdens on other agents. (For more about this and how to use this dysfunctional behavior to model the behavior during a transition between social norms, see Saari 2008, Chap. 2.)

Finally, return to the Sect. 3.3.1 question as to whether the Borda Count should be criticized because it need not elect a Condorcet winner, or whether the Condorcet winner should be criticized because it need not be the Borda winner. As shown in Theorem 3, all possible differences between Borda and paired comparison tallies, between the Borda and Condorcet winners, are caused by a profile's $\mathcal{RWC}$ portion. In turn, it follows from Sect. 3.4.2 that if the Condorcet winner is not the Borda winner, the reason is that in selecting the Condorcet winner, the rule is reflecting the wishes of *nonexistent voters who have cyclic preferences.* Because it seems wiser to side with rules that reflect the preferences of actual voters, rather than nonexistent ones, the appropriate criticism clearly is against the Condorcet winner;

> *a fault suffered by paired comparisons and the Condorcet winner is that they need not elect the Borda winner.*

---

[3] Readers interested in understanding why, say, the plurality vote does not qualify for Theorem 5 can find answers in (Saari 2008, Chap. 4). It is shown there why all positional methods, other than the Borda Count, can completely ignore properties of paired comparisons.

4 General Paired Comparisons

It remains to explain the source of problems afflicting methods such as AHP and to relax the condition that paired comparisons rely on profiles with complete transitive rankings. This is done by ignoring the source of paired comparison values and concentrating on their consequences. In doing so, a selection of new results developed in Saari (2010) are discussed.

4.1 A Coordinate System

With $N \geq 3$ options or alternatives, there are $\binom{N}{2} = \frac{N(N-1)}{2}$ pairs. In representing each pair, the idea is to generalize the "tally-difference" $\tau(X, Y)$ notation. Just as $\tau(X, Y) = -\tau(Y, X)$, the value assigned to a pair $\{A_i, A_j\}$ of alternatives is $d_{i,j}$ where

$$d_{i,j} = -d_{j,i}. \tag{21}$$

One interpretation for these values is $d_{i,j} = \tau(X_i, X_j)$, but there are many others. Thus, the $d_{i,j}$ term is left undefined until specified in examples.

Because of Eq. 21, it suffices to consider only $d_{i,j}$ terms where $i<j$. When $d_{j,i}$ $j > i$, is needed, replace $d_{j,i}$ with $-d_{i,j}$. Thus all relevant terms can be assembled in a vector

$$\mathbf{d} = (d_{1,2}, d_{1,3}, , d_{1,N}; d_{2,3}, , d_{2,N}; , d_{N-1,N}) \in \mathbb{R}^{\binom{\mathbb{N}}{2}}, \tag{22}$$

where the semicolons indicate a change in the first subscript.

4.1.1 Strongly Transitive Terms

Just as "strong transitivity" was defined in terms of the τ's in Eq. 12, vector $\mathbf{d}$ is said to be strongly transitive if each triplet $\{i, j, k\}$ satisfies

$$d_{i,j} + d_{j,k} = d_{i,k}. \tag{23}$$

Theorem 6. (Saari 2010) The strongly transitive choices of $\mathbf{d} \in \mathbb{R}^{\binom{\mathbb{N}}{2}}$ define a linear subspace of dimension $N - 1$. This subspace is denoted by $\mathcal{ST}_N$.

The linearity of $\mathcal{ST}_N$ and its dimension follow from Eq. 23. To see the dimension assertion, notice that all components of $\mathbf{d} \in \mathcal{ST}_N$ are uniquely determined from the $d_{1,2}, d_{2,3}, , d_{N-1,N}$ values. The $d_{1,3}$ value, for instance, follows from $d_{1,2} + d_{2,3}$, the $d_{1,4}$ value is given by $d_{1,3} + d_{3,4} = [d_{1,2} + d_{2,3}] + d_{3,4}$, and so forth. The proof of Theorem 6 follows in a similar simple manner.

This dimension assertion introduces a mystery because it differs from the Theorem 2 assertion that $\mathcal{ST}_N$ had the larger dimension of $(2N-1)\frac{(N-1)!}{2}$. Already with $N=3$, the strongly transitive profiles from Theorem 2 live in a *five-dimensional space* while Theorem 6 allows only *two-dimensions*. The difference becomes more dramatic with larger values of N; e.g., with $N=5$, the $\mathcal{ST}$ profiles reside in an 108-dimensional space, while Theorem 6 choices are confined to a meager four-dimensional space— a shocking difference of 104 dimensions.

The explanation is that Theorem 2 describes subspaces of *profiles*, while Theorem 6 describes subspaces of outcomes. In fact, most $\mathcal{ST}$ profiles create paired comparison ties. The two-person profile $\{A \succ B \succ C,\ C \succ B \succ A\}$, for instance, is $\mathcal{ST}$ because $\tau(A,B)=\tau(B,C)=\tau(A,C)=0$; this profile is mapped to the zero vector in Theorem 6. Indeed, this "zero outcome" is the fate for a subspace of $\mathcal{ST}$ profiles with the dominating dimension of $\left[(2N-1)\frac{(N-1)!}{2}-(N-1)\right]$. (This huge subspace is further divided to extract new kinds of "ranking wheel symmetries" that are needed to analyze the triplets and higher order clustering that were mentioned in the concluding sentence of Sect. 1.)

What simplifies using the $\mathcal{ST}$ subspace is that it has a convenient basis.

Definition 3. For each $i=1,,N$, let vector $\mathbf{B}_i \in \mathbb{R}^{\binom{\mathbb{N}}{\neq}}$ be where each $d_{i,j}=1$ for $j\neq i$, $j=1,,N$, and each $d_{k,j}=0$ if $k,j\neq i$. $\mathbf{B}_i$ is called the "A_i-basic vector."

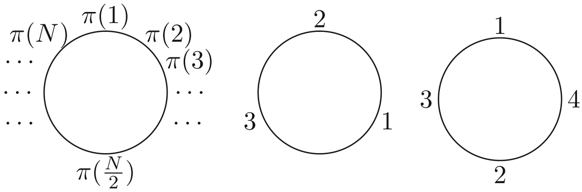
As examples with $N=4$, $\mathbf{B}_1=(1,1,1;0,0;0)$, and $\mathbf{B}_2=(-1,0,0;1,1;0)$. The "$-1$" in the $d_{1,2}$ position of $\mathbf{B}_2$ is because Definition 3 requires $d_{2,1}=1$; Eq. 21 converts this value to $d_{1,2}=-1$. Similarly $\mathbf{B}_3=(0,-1,0;0,0;1)$; these three vectors serve as a basis. Thus a vector in $\mathcal{ST}_4$ can be expressed as

$$\sum_{j=1}^{3}\alpha_j\mathbf{B}_j=(\alpha_1-\alpha_2,\alpha_1-\alpha_3,\alpha_1;\alpha_2-\alpha_3,\alpha_2;\alpha_3).$$

That any such choice satisfies the $\mathcal{ST}$ condition of Eq. 23 reduces to simple arithmetic. It must be shown, for instance, that $d_{1,2}+d_{2,3}=(\alpha_1-\alpha_2)+(\alpha_2-\alpha_3)=\alpha_1-\alpha_3$ equals $d_{1,3}=\alpha_1-\alpha_3$, which is immediate. In much the same way, Theorem 7 is proved.

Theorem 7. (Saari 2010) For $N\geq 3$, any subset of $(N-1)$ vectors from $\{\mathbf{B}_i\}_{i=1}^{N}$ spans $\mathcal{ST}_N$.

Now that the $\mathcal{ST}$ components of vectors is defined, the next step is to identify the orthogonal complement. This development introduces $\mathcal{RWC}$ terms.



(a) General case **(b)** $N = 3$ **(c)** $N = 4$

Fig. 2 Ranking wheel arrangements of data

4.1.2 Ranking Wheel Configurations

As it must be expected from Theorem 2, the rest of this $\mathbb{R}^{\binom{\mathbb{N}}{\nvdash}}$ space of pairwise outcomes is characterized by $\mathcal{RWC}$ behavior. The following Definition 4 describes a behavior that parallels the Fig. 1 definition of a $\mathcal{RWC}$.

Definition 4. Let π be a permutation of the indices $1, 2, , N$ as listed in the cyclic fashion $(\pi(1), \pi(2), , \pi(N))$ around a circle. Define $\mathbf{C}_\pi \in \mathbb{R}^{\binom{\mathbb{N}}{\nvdash}}$ as follows: If j immediately follows i in a clockwise direction, then $d_{i,j} = 1$; if j immediately precedes i, then $d_{i,j} = -1$. Otherwise $d_{i,j} = 0$. Vector $\mathbf{C}_\pi$ is the "$\mathcal{RWC}$ direction defined by π".

The definition is depicted in Fig. 2a where the index numbers for permutation π are listed around the ranking wheel as in Fig. 1. The $N = 3$ permutation $\pi = (2, 1, 3)$ is represented in Fig. 2b. To compute $\mathbf{C}_\pi = (d_{1,2}, d_{1,3}; d_{2,3})$, because 2 precedes 1 on the circle, $d_{1,2} = -1$, because 3 immediately follows 1, $d_{1,3} = 1$, and because 3 immediately precedes 2, $d_{2,3} = -1$. These terms define $\mathbf{C}_{(2,1,3)} = (-1, 1; -1)$.

Vector $\mathbf{C}_{(2,1,3)}$ violates the $\mathcal{ST}$ (Eq. 23) condition. Conflicting with $d_{1,2} + d_{2,3} = d_{1,3}$, the $\mathbf{C}_{(2,1,3)}$ entries are $d_{1,2} = d_{2,3} = -1$, where, rather than the required $d_{1,3} = -2$, the other extreme of $d_{1,3} = 1$ occurs. Expressed with rankings, these $d_{i,j}$ terms represent the cycle $A_2 \succ A_1, A_1 \succ A_3$, $A_3 \succ A_2$ where (as to be expected from $\mathcal{RWC}$) each pair has the same $d_{i,j}$ difference. What sharpens the connection is when this $\mathcal{RWC}$ has a profile representation, it follows that $\tau(A_2, A_1) = -\tau(A_1, A_2), \tau(A_1, A_3)$, and $\tau(A_3, A_1) = -\tau(A_1, A_3)$ all equal unity, which agrees with $\mathbf{C}_{(2,1,3)} = (-1, 1; -1)$.

A direct computation also proves that the $\mathcal{RWC}$ direction $\mathbf{C}_{(2,1,3)}$ is orthogonal to $\mathbf{B}_1 = (1, 1; 0)$ and $\mathbf{B}_2 = (-1, 0; 1)$, so it is orthogonal to $\mathcal{ST}_3$. As these $\mathbf{C}_\pi$ vectors span the space orthogonal to $\mathcal{ST}_N$, an assertion consistent with Theorem 2 follows.

Theorem 8. For any $N \geq 3$ and permutation π of the indices, vector $\mathbf{C}_\pi$ is orthogonal to $\mathcal{ST}_N$. These $\mathbf{C}_\pi$ vectors span a space denoted by $\mathcal{RWC}_N$, which is the $[\binom{N}{2} - (N-1)] = \binom{N-1}{2}$-dimensional subspace of vectors that are orthogonal to $\mathcal{ST}_N$.

As soon as $(N-1) \geq \binom{N-1}{2}$, or $N \geq 4$, the dimension of $\mathcal{RWC}$ components is as large as, or greater than, the dimension of $\mathcal{ST}$ data components. Therefore, *as soon as $N \geq 5$, there are more data dimensions of the trouble-generating $\mathcal{RWC}$ terms than of $\mathcal{ST}$ directions that have consistent outcomes.* A worrisome corollary is that data must be anticipated to include $\mathcal{RWC}$ cyclic outcomes. As before, paired comparison rules experience difficulties because they are almost always affected by data with $\mathcal{RWC}$ components.

4.1.3 General Outcomes

Theorems 7 and 8 provide a tool analogous to Theorem 2. That is, for any rule depending on paired comparisons, whether or not the outcomes are based on complete transitive profiles, the properties of the rule can be analyzed. An example doing so for nonparametric statistics is in Bargagliotti and Saari (2010), while Saari and Sieberg (2004) identifies difficulties that arise in engineering.

In fact, this structure can be used to prove Theorem 1. This is because, while Theorem 1 is described in terms of profiles, the conclusion about strongly transitive outcomes is based on how $\tau(X_i, X_j)$ values interact, and these $\tau(X_i, X_j)$ terms are captured here by the $d_{i,j}$ terms. The above shows that $\mathbf{d}$ choices that are orthogonal to the cyclic behaving $\mathcal{RWC}$ components satisfy the strongly transitive condition; this is the Theorem 1 assertion.

Theorem 7 also simplifies discovering extensions to Theorem 1. An example follows:

Theorem 9. For $N \geq 3$, dropping an option from a $\mathcal{ST}$ data set (or profile) creates a $(N-1)$-option $\mathcal{ST}$ data set (or profile). However, dropping an option from a $\mathcal{RWC}$ data set (or from a $\mathcal{RWC}$ component of a profile) creates a $(N-1)$-option $\mathcal{RWC}$ data set (or profile) *along with* a new $(N-1)$-option $\mathcal{ST}$ component.

Armed with the basis vectors, Theorem 9 is easy to prove. With N alternatives and $\mathbf{B}_1 = (1, 1, , 1; 0, ; 0)$, for instance, dropping any alternative merely removes one of the "1's" and a block of the zeros creating a basis vector for $(N-1)$-options. A similar argument holds for $\mathbf{C}_\pi$ vectors. Illustrating with $\mathbf{C}_{1,4,2,3}$ depicted in Fig. 2c, it is given by

$$\mathbf{C}_{1,4,2,3} = (0, -1, 1; 1, -1; 0).$$

Dropping option 4 creates

$$\mathbf{C} = (0, -1; 1) = \frac{2}{3}(1, -1; 1) - \frac{1}{3}(1, 1; 0) + \frac{1}{3}(-1, 0; 1) = \frac{1}{3}[2\mathbf{C}_{1,2,3} + \mathbf{B}_2 - \mathbf{B}_1].$$

An insightful way to see this conclusion is to use a $\mathcal{RWC}$ profile and drop an alternative; e.g., with the Eq. 5 profile, dropping F creates a $\mathcal{RWC}$ with initial ranking $A \succ B \succ C \succ D \succ E$ *along with the extra ranking of* $A \succ B \succ C \succ D \succ E$.

While simple to prove, Theorem 9 solves many long-standing mysteries; it asserts that if a profile, or a data set, has a reasonably sized $\mathcal{RWC}$ component, then dropping options can reverse the ranking. It is interesting that changing the choice of the dropped option changes the choice of the extra ranking *in a cyclic manner*. Illustrating with $N = 4$ and the $\mathcal{RWC}$ generated by $A \succ B \succ C \succ D$, a computation shows that

- when D is not being considered, the extra ranking is $A \succ B \succ C$;
- when A is not being considered, the extra ranking is $B \succ C \succ D$;
- when B is not being considered, the extra ranking is $C \succ D \succ A$;
- when C is not being considered, the extra ranking is $D \succ A \succ B$

where each candidate is in each position in precisely one choice. Indeed, this cycle of extra rankings is what creates the 2:1 cycle described in the introduction of Sect. 3 and it causes the cycle of plurality votes when considering the six quintuplets in Eq. 2. As another consequence, this cycle of extra rankings coming from $\mathcal{RWC}$ is the only and complete source of inconsistencies in Borda rankings. (So if the Borda ranking of N alternatives disagrees with the Borda ranking for any $k < N$ alternatives, it is due to a $\mathcal{RWC}$ component in the profile; the effect of this component on Borda rankings is as described in Theorem 9.)

In fact, because Theorem 9 is the total explanation for changes in rankings when analyzing paired comparisons, it is reasonable to wonder if it explains the AHP problem described in Sect. 2.2 with Eq. 3. It does; to demonstrate this fact, basic structures need to be extended so that versions of Theorems 7, 8, and 9 are applicable. This is indicated in the next section.

4.2 Analyzing AHP Complexities

Recall from Sect. 2.2 that $a_{i,j} = \frac{1}{a_{j,i}}$, which allows the AHP data to be represented in a fashion similar to Eq. 22 as

$$\mathbf{a} = (a_{1,2}, a_{1,3}, , a_{1,N}; a_{2,3}, , a_{2,N}; ; a_{N-1,N}) \in \mathbb{R}_+^{\binom{\mathbb{N}}{\not{\mathbb{Z}}}}$$

where the "+" subscript means that all entries are positive. With AHP, a matrix is "consistent" if for each triplet {i, j, k}, the equality

$$a_{i,j} a_{j,k} = a_{i,k} \tag{24}$$

is satisfied. A delightful feature of consistent AHP matrices (e.g., Saaty 1980) is that the $a_{i,j}$ values satisfy the strong connection with eigenvector $\mathbf{w}$ given by

$$a_{i,j} = \frac{w_i}{w_j} \quad \textit{for all } i, j. \tag{25}$$

Consistency is such an important feature that the surface of vectors **a** that satisfy this condition deserves its own name.

Definition 5. For $N \geq 3$, let $\mathbb{C}_N = \{\mathbf{a} \in \mathbb{R}_+^{\binom{N}{2}} \mid \mathbf{a}$ *satisfies Eq.*24$\}$. This space $\mathbb{C}_\mathbb{N}$ is called the "consistency manifold of $\mathbb{R}_+^{\binom{\mathbb{N}}{\nvDash}}$."

It follows from the consistency equation (Eq. 24) that $\mathbb{C}_\mathbb{N}$ is a smooth $(N-1)$-dimensional manifold of $\mathbb{R}_+^{\binom{\mathbb{N}}{\nvDash}}$. In fact, holding $a_{i,k}$ fixed defines a hyperbola, which shows that the geometry of $\mathbb{C}_\mathbb{N}$ has a hyperbolic component. While the importance of this surface is not obvious at this point, its value will quickly become apparent.

4.2.1 Connecting Multiplicative with Additive Structures

To establish a connection between Sect. 4.1 and the AHP structure, define the mapping

$$a_{i,j} = e^{d_{i,j}}, \quad or \quad d_{i,j} = \ln(a_{i,j}) \, for \; all \; i,j. \tag{26}$$

This Eq. 26 mapping maps $\mathbb{R}^{\binom{\mathbb{N}}{\nvDash}}$ onto $\mathbb{R}_+^{\binom{\mathbb{N}}{\nvDash}}$ in a one-to-one smooth manner. The pragmatic effect of this mapping is to transfer the additive structure developed in Sect. 4.1 to the multiplicative setting of AHP. In turn, the $\mathcal{ST}_N$ and $\mathcal{RWC}_N$ subspaces can be expected to have counterparts that will explain AHP problems.

Because $\ln(a_{i,j}a_{j,k}) = \ln(a_{i,j}) + \ln(a_{j,k})$, Eq. 24 is equivalent to

$$\ln(a_{i,j}) + \ln(a_{j,k}) = \ln(a_{i,k}), \quad \text{or, with Eq. 26,} \quad d_{i,j} + d_{j,k} = d_{i,k}.$$

Thus, consistency and strongly transitive are equivalent; the surfaces $\mathbb{C}_\mathbb{N}$ and $\mathcal{ST}_N$ are mapped onto each other. Of particular interest, this connection means that "consistency" plays an identical role for AHP as strongly transitive structures play for additive settings.

This comment allows us to completely analyze a class of AHP problems. For instance, it follows from $\mathcal{ST}$ properties (Theorem 9) that if an AHP matrix is consistent, then should any option be dropped, the resulting matrix remains consistent and relative weights assigned to alternatives remain unchanged. Thus all Eq. 3 kinds of problems must be due to something that is equivalent to $\mathcal{RWC}$ terms.

To illustrate, the Eq. 3 example defines the vector $\mathbf{a} = (6, \frac{3}{4}; 8)$, which, when mapped by Eq. 26 to the linear structure becomes $\mathbf{d} = (\ln(6), \ln(\frac{3}{4}); \ln(8))$. This system, where $d_{1,2} = \ln(6)$, $d_{1,3} = \ln(\frac{3}{4})$, and $d_{2,3} = \ln(8)$ is far from satisfying the Eq. 23 equation for $\mathcal{ST}$ status. To find the trouble-making $\mathcal{RWC}$ component in this data set, find the value of $\ln(x)$ with $\mathbf{C}_{1,3,2}$ (see Sect. 4.1.2) so that

$$\left(\ln(6) - \ln(x),\ \ln\left(\frac{3}{4}\right) + \ln(x);\ \ln(8) - \ln(x)\right) = \left(\ln\left(\frac{6}{x}\right),\ \ln\left(\frac{3x}{4}\right);\ \ln\left(\frac{8}{x}\right)\right)$$

does satisfy Eq. 12. Using the equivalent Eq. 24, the value of x must satisfy

$$\frac{6}{x} \times \frac{8}{x} = \frac{3x}{4} \quad or \quad x^3 = 64; \quad x = 4.$$

Thus the consistent part of the Eq. 3 matrix is given by $\mathbf{a} = (\frac{3}{2}, 3; 2)$, which defines the consistent matrix

$$\begin{pmatrix} 1 & \frac{3}{2} & 3 \\ \frac{2}{3} & 1 & 2 \\ \frac{1}{3} & \frac{1}{2} & 1 \end{pmatrix}$$

with eigenvector $\mathbf{w} = (\frac{1}{2}, \frac{1}{3}, \frac{1}{6})$ that is the same as for Eq. 3. As before, A_1 has three times the intensity assigned to the bottom ranked option A_3. But with this *consistent* matrix, dropping option A_2 creates matrix $\begin{pmatrix} 1 & 3 \\ \frac{1}{3} & 1 \end{pmatrix}$ with eigenvector $(\frac{3}{4}, \frac{1}{4})$ where A_1 *retains its 3:1 advantage over* A_3. In other words, the Sect. 2.2 paradoxical behavior for AHP is strictly caused by the $\mathcal{RWC}$ component of the Eq. 3 matrix! The negative aspects of $\mathcal{RWC}$ are manifested again! This suggests that a way to refine AHP outcomes while eliminating its paradoxical outcomes is to develop a way to eliminate the insidious $\mathcal{RWC}$ cyclic effects. A simple way to do so is described in Saari (2010).

To conclude, earlier I asserted that examples can be created where the difference between the top and bottom ranked AHP alternative has any desired multiple difference even though the paired comparison between these options reverses the inequality. To do so, let the multiple between A_1 and A_3 be M. Select any multiple for the A_1 and A_2 difference, say 2. A consistent setting requires $a_{1,2} = 2$, $a_{1,3} = M$, which means from Eq. 23 that $a_{1,2}a_{2,3} = 2a_{2,3} = a_{1,3} = M$, or $a_{2,3} = \frac{M}{2}$. Thus the corresponding consistent matrix is

$$\begin{pmatrix} 1 & 2 & M \\ \frac{1}{4M} & 1 & \frac{M}{2} \\ \frac{1}{M} & \frac{2}{M} & 1 \end{pmatrix}$$

with (Eq. 25) eigenvector $\mathbf{w} = \left(\frac{M}{M+\frac{M}{2}+1}, \frac{M}{3M+2}, \frac{1}{M+\frac{M}{2}+1}\right)$ and the desired $A_1 \succ A_2 \succ A_3$ ranking; the intensity weights have A_1 being M times better than A_3.

According to the above (particularly Theorem 9), the desired conclusion requires adding a strong $\mathcal{RWC}$ effect to the data. By using Eq. 26 and $\mathbf{C}_{1,2,3} = (1, -1; 1)$, this requires finding a x value so that

$$d_{1,2} = \ln(2) + \ln(x);\ d_{1,3} = \ln(M) - \ln(x) < 0;\ d_{2,3} = \ln\left(\frac{M}{2}\right) + \ln(x)).$$

The only constraint (to ensure that $A_3 \succ A_1$ in the paired comparison) is that $d_{1,3} = \ln(\frac{M}{x}) < 0$. A simple choice is $x = 2M$, which defines $\mathbf{a} = (4M, \frac{1}{2}; M^2)$ and the matrix

$$\begin{pmatrix} 1 & 4M & \frac{1}{2} \\ \frac{1}{4M} & 1 & M^2 \\ 2 & \frac{1}{M^2} & 1 \end{pmatrix},$$

which has the same eigenvector $\mathbf{w} = \left(\frac{M}{M+\frac{M}{2}+1}, \frac{M}{3M+2}, \frac{1}{M+\frac{M}{2}+1}\right)$ asserting that A_1 is M times better than A_3.[4] But when A_2 is dropped to compare the extremes with

$$\begin{pmatrix} 1 & \frac{1}{2} \\ 2 & 1 \end{pmatrix},$$

the eigenvector $\left(\frac{1}{3}, \frac{2}{3}\right)$ now has A_3 *being twice as good as* A_1.

5 Summary

The world of paired comparisons has been full of mysteries. But as demonstrated above, *all of these problems*, whether manifested in Arrow's theorem, Sen's assertion, problems with the Condorcet winner, the Borda Count, AHP, statistical properties, engineering decisions, economic analysis, or with anything, are caused by the $\mathcal{RWC}$ portion of the data. This simple ranking wheel has the ability to explain a variety of previously troubling problems and to indicate how to find positive resolutions. A book is being prepared to show how ranking wheel configurations can relate, explain, and extend paired comparison results including those from spatial voting (such as the existence of a core).

Acknowledgments My thanks to M. Holler and K. Sieberg for several helpful comments about earlier versions of this paper. This research was supported by NSF grants DMS-0631362 and CMMI-1016785. This paper is an extended version of Saari (2011).

References

Arrow, K. (1951). Social choice and individual values (2nd ed. 1963). New York: Wiley.

Bargagliotti, A., & Saari, D. G. (2010). Symmetry of nonparametric statistical tests on three samples. *Journal of Mathematics and Statistics*, *6*(4), 395–408.

[4] As shown in Saari (2010), adding a *single* $\mathcal{RWC}$ to a consistent matrix keeps the same eigenvector.

Condorcet, M. Éssai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. Paris.

Kemeny, J. (1959). Mathematics without numbers. *Daedalus*, *88*, 571–591.

Nakamura, K. (1978). The voters in a simple game with ordinal preferences. *International Journal of Game Theory*, *8*, 55–61.

Saari, D. G. (1995). *Basic geometry of voting*. New York: Springer.

Saari, D. G. (2000). Mathematical structure of voting paradoxes 1; pairwise vote. *Economic Theory*, *15*, 1–53.

Saari, D. G. (2008). *Disposing dictators, demystifying voting paradoxes*. Cambridge: Cambridge University Press.

Saari, D. G. (2010). Geometry of voting; Chap. 27. In K. Arrow, A. Sen & K. Suzumura (Eds.), *Handbook of social choice & welfare* (pp. 897–946). San Diego: Elsevier.

Saari, D. G. (2010) A new way to analyze paired comparison rules, IMBS Discussion papers.

Saari, D. G. (2011). Ranking wheels and decision cycles. *Homo Oeconomicus*, *28*, 233–263.

Saari, D. G., & Sieberg, K. (2004). Are part wise comparisons reliable? *Research in Engineering Design*, *15*, 62–71.

Saaty, T. (1980). *The analytic hierarchy process: Planning, priority setting, resource allocation*. New York: Mcgraw-Hill.

Sen, A. (1970a). The impossibility of a Paretian liberal. *Journal Political Economy*, *78*, 152–157.

Sen, A. (1970b). *Collective choice and social welfare*. San Francisco: Holden Day.

A Geometric Approach to Paradoxes of Majority Voting: From Anscombe's Paradox to the Discursive Dilemma with Saari and Nurmi

Daniel Eckert and Christian Klamler

1 Introduction

In the last thirty years, there have been several attempts to generalize the Arrovian framework of preference aggregation (e.g. Rubinstein and Fishburn 1986 or Wilson 1975). This literature on abstract aggregation has been considerably stimulated by the growing interest in problems of judgment aggregation. The problem of judgment aggregation consists in aggregating individual judgments on an agenda of logically interconnected propositions into a collective set of judgments on these propositions (see List and Puppe 2009 for a survey).

As an example of a paradox in judgment aggregation, consider a variant of the so-called discursive dilemma, in which a committee of three recruitment officers in a firm has to decide whether a job applicant should be hired or not. There is a written test and an oral interview and each of them is advised to recommend hiring the applicant if and only if the applicant passes the written test and gives a satisfiable interview. Table 1 shows the judgments of the officers and their majority decisions.

Based on the majority of the individual decisions, the job applicant will not be hired as a majority does not find her acceptable. However, a majority finds the written test as well as the interview acceptable.

An earlier version of this chapter has been published in Homo Oecomicus 26(3/4): Essays in Honor of Hannu Nurmi: Volume I, edited by Manfred J. Holler and Mika Widgren, 2009.

D. Eckert (✉) · C. Klamler
Institute of Public Economics, University of Graz, 8010 Graz, Austria
e-mail: daniel.eckert@uni-graz.at

C. Klamler
e-mail: christian.klamler@uni-graz.at

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_33, © Springer-Verlag Berlin Heidelberg 2013

Table 1 Discursive dilemma

Officer	Written test	Oral interview	Decision
Officer 1	1	1	1
Officer 2	1	0	0
Officer 3	0	1	0
Majority outcome	1	1	0

Problems of judgment aggregation are structurally similar to paradoxes and problems in social choice theory like the Condorcet paradox and Arrow's general possibility theorem, but also related to paradoxes of compound majorities like the Anscombe or Ostrogorski paradoxes, both nicely analysed by Nurmi (1997) (see also Nurmi 1987, 1999). Contemporarily, Saari (1995) has developed and popularized a geometric approach to Arrovian social choice theory. His approach has greatly helped to understand what drives many of the impossibility results and paradoxes in social choice theory.

In a similar vein falls Nurmi's (2004) distance-based approach to aggregation problems, something that can be seen as one of the earliest attempts to apply a geometric approach to abstract aggregation problems. For example Meskanen and Nurmi (2006) show that all of the aggregation rules devised to overcome the problem of majority cycles (e.g. Copeland rule, Kemeny rule) can be characterized by a distance measure and a certain goal state.

In this chapter we develop—in Saari's style—a geometric approach to abstract aggregation theory starting from a paradox intensively investigated by Nurmi and extending this framework to typical paradoxes in judgment aggregation.

Our approach focuses on and will be exhaustive for aggregation problems that can be represented in the three-dimensional hypercube. While this is the smallest dimension in which interesting aggregation problems can be formulated and be particularly illuminating for problems that naturally fall into this framework, we have to give a warning that most of our results are not easily extendable to more than three dimensions.

A major difference of judgment aggregation to social choice theory lies in the representation of the information involved. While binary relations over a set of alternatives are a canonical representation of preferences, a natural representation of judgments are binary valuations over a set of propositions, where the logical interconnections between these propositions determine the set of admissible valuations. For example the agenda of the famous discursive dilemma $\{p, q, p \wedge q\}$ is associated the set of admissible, i.e. logically consistent valuations $\{(0, 0, 0), (1, 0, 0), (0, 1, 0), (1, 1, 1)\}$, where a 1(0) denotes a proposition to be believed (not believed).

The chapter is structured as follows: In Section 2 we introduce the formal framework. Section 3 discusses paradoxes of majority voting. We will use Saari's representation cubes to provide a unified geometric representation of profiles and majority rule outcomes and introduce Saari's idea of a profile decomposition. In this framework we will provide a characterization of profiles leading to the Anscombe paradox. Section 4 applies the same tools to judgment aggregation. In particular we show what drives the logical inconsistency of majority outcomes and how this can be avoided with the help of restrictions on the distribution of individual valuations, i.e. give a kind of generalized domain restriction. This leads us to the determination of the likelihood of inadmissible outcomes under majority rule for different agendas in Section 5. In Section 6, we apply our approach to illuminate current results on distance-based judgment aggregation. Finally, Section 7 concludes the chapter.

2 Abstract Aggregation Theory and Majority Voting

In the binary framework of abstract aggregation theory individual vectors of yes/no or true/false valuations $v = (v^1, v^2, \ldots, v^{|J|}) \in \{0,1\}^{|J|}$ from a set $X \subseteq \{0,1\}^{|J|}$ of admissible valuations over a set J of issues (the agenda) are aggregated into a collective valuation. (In a slight abuse of notation we will use the term valuation both for the binary valuation of a single issue as for vectors of binary valuations.)

Such an issue $j \in J$ might be the pairwise comparison between two alternatives in preference aggregation or a proposition on which a judgment needs to be made. Typically, the interconnections between the issues limit the set of admissible valuations. In judgment aggregation a valuation $v = (v^1, v^2, \ldots, v^{|J|}) \in X \subseteq \{0,1\}^{|J|}$, represents an individuals' beliefs, where $v^j = 1$ means that proposition j is believed and X denotes the set of all admissible (logically consistent) valuations (see Dokow and Holzman 2010).

Given a set N of individuals, a profile of individual valuations is then a mapping $p : N \rightarrow \{0,1\}^{|J|}$ which assigns to each individual a vector of binary valuations. A desirable property of an aggregation rule, stronger than non-dictatorship, is of course anonymity, which requires that the same collective valuation be assigned to any permutation of the set of individuals.

If anonymity is assumed, a profile of individual valuations can be represented by a vector $\mathbf{p} = (p_1, \ldots, p_{|X|}) \in [0,1]^{|X|}$ with $\sum_k p_k = 1$, which associates with every admissible valuation $v_k \in X$ the share p_k of individuals with this valuation. Such an anonymous representation of profiles is particularly appropriate for the analysis of majority voting, where anonymity is typically assumed.

Geometrically any binary valuation is a vertex of the $|J|$-dimensional hypercube and, more interestingly, any anonymous profile $\mathbf{p} \in [0,1]^{|X|}$ can be given a lower-dimensional representation by a point $x(\mathbf{p}) \in [0,1]^{|J|}$ in the $|J|$-dimensional 0/1-polytope, i.e. the convex hull of the hypercube $\{0,1\}^{|J|}$, where for each component $j \in J$, $x^j(\mathbf{p}) = \sum_{k \in \{1,\ldots,|X|\}} p_k v_k^j$ denotes the average support for issue j. Thus the $|J|$-dimensional 0/1-polytope will be referred to as the representation polytope of the profiles.

An abstract anonymous aggregation rule is a mapping f that associates with every anonymous profile $\mathbf{p} = (p_1, p_2, \ldots, p_{|X|}) \in [0,1]^{|X|}$ a valuation $v = f(x(\mathbf{p})) \in \{0,1\}^{|J|}$.

We will write $v(\mathbf{p})$ for $f(x(\mathbf{p}))$ and identify by $v^j(\mathbf{p})$ the jth component of $v(\mathbf{p})$ under the given aggregation rule.

In this framework majority voting on issues (or majority voting for short) is defined as follows:

Definition 1 *For any issue* $j \in J$ and any profile $\mathbf{p} \in [0,1]^{|X|}$, ${}^M v^j(\mathbf{p}) \in \{0,1\}$ is the outcome of majority voting on issue j if $v^j(\mathbf{p}) = 1 \Leftrightarrow x^j(\mathbf{p}) > 0.5$.

This representation immediately provides majority with a wellknown metric rationalization in terms of the Hamming distance between binary vectors. (For any two binary vectors $v, v' \in \{0,1\}^{|J|}$, the Hamming distance $d_H(v, v')$ is the number of components in which these two vectors differ.)

Proposition 1 (Brams et al. 2004) *For any profile* $\mathbf{p} \in [0,1]^{|X|}$, *the valuation* ${}^M v(\mathbf{p}) \in \{0,1\}^{|J|}$ *is the majority outcome if and only if it minimizes the sum of Hamming distances weighted by the population shares, or formally*,

$$ {}^M v(\mathbf{p}) = \underset{v \in \{0,1\}^{|J|}}{\arg\min} \sum_{k=1}^{|X|} p_k d_H(v_k, v). $$

Thus, whenever the sum of Hamming distances can be interpreted as an appropriate measure of social disutility, majority voting can be justified by its minimization.

Observe however that nothing in this characterisation prevents the majority outcome ${}^M v(\mathbf{p}) \in \{0,1\}^{|J|}$ from being an inadmissible valuation, i.e. that ${}^M v(\mathbf{p}) \in \{0,1\}^{|J|} \backslash X$.

In the hypercube, a more natural metric representation of majority voting can be given in terms of the euclidean distance d_E.

Proposition 2 *For any profile* $\mathbf{p} \in [0,1]^{|X|}$, *the valuation* ${}^M v(\mathbf{p}) \in \{0,1\}^{|J|}$ *is the majority outcome if and only if it minimizes the euclidean distance between the corresponding vertex and the point* $x(\mathbf{p})$ *in the representation polytope, or formally*

$$ {}^M v(\mathbf{p}) = \underset{v \in \{0,1\}^{|J|}}{\arg\min}\, d_E(x(\mathbf{p}), v). $$

Conversely, the set of all profiles for a given majority outcome $v \in \{0,1\}^{|J|}$ *defines a subcube of* $[0,1]^{|J|}, P_v = [|v^j - 0.5|]^{|J|}$, *which is the set of all profiles for which* v *is the majority outcome. Such a subcube will be called the majority subcube of* v (*or simply* v-*subcube*) *and can be seen in* Fig. 1 *for vertex* $(1,0,1)$.

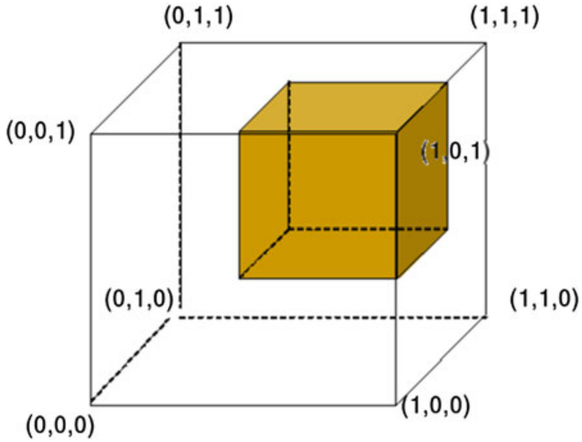


Fig. 1 Majority subcube

3 The Anscombe Paradox and the Irrationality of a Metric Rationalization

Because majority voting on issues has a metric rationalization in terms of distance minimization, it is quite disturbing that the majority outcome need not be the one that minimizes the distance for the majority of individuals, as the Anscombe paradox shows. In other words the Anscombe paradox states that a majority of the voters can be on the loosing side on a majority of issues. Formally, the Anscombe paradox can be defined in the following way:

Definition 2 A profile $\mathbf{p} = (p_1, \ldots, p_{|X|}) \in [0,1]^{|X|}$ exhibits the Anscombe paradox if

$$\sum_{k \in \{1,\ldots,|X|\}: d_H(v_k, {}^M v(\mathbf{p})) > \frac{|J|}{2}} p_k > \frac{1}{2}$$

Indeed, it is the particular distribution of individual valuations that leads to the paradox. To analyse this and further paradoxes in later sections, we will numerate the vertices of the three-dimensional hypercube as listed in Table 2.

Now, consider the profile $\mathbf{p} = (\frac{2}{5}, 0, 0, 0, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, 0)$, which specifies exactly an Anscombe paradox situation. It is easily observed, that $x(\mathbf{p}) = (\frac{2}{5}, \frac{2}{5}, \frac{2}{5})$ and hence the majority outcome is ${}^M v(\mathbf{p}) = (0, 0, 0)$.

Are we able to specify profiles that lead to an Anscombe type or other paradoxical majority outcome? Saari (2008) identifies what he calls "Condorcet portions" as the driving part of paradoxes of preference aggregation.[1] In our three-

[1] A "Condorcet portion" is a multiple of the set of individuals that has the following preferences over 3 alternatives a, b, c: $a \succ_1 b \succ_1 c$, $c \succ_2 a \succ_2 b$, $b \succ_3 c \succ_3 a$ leading to the the majority cycle $a \succ b \succ c \succ a$.

Table 2 Valuations in three-dimensional hypercube

Valuation		Valuation	
v_1	(0, 0, 0)	v_5	(1, 1, 0)
v_2	(1, 0, 0)	v_6	(1, 0, 1)
v_3	(0, 1, 0)	v_7	(0, 1, 1)
v_4	(0, 0, 1)	v_8	(1, 1, 1)

dimensional setting for abstract aggregation problems, we can consider such portions as triples of valuations that have a common neighbor, i.e. a valuation that differs from each of the three valuations in exactly one issue.[2] Given that, we can now easily specify for every vertex in the hypercube its triple of neighbors. For example for v_5 the corresponding triple of neighbors is (v_2, v_3, v_8). Table 3 indicates the triples for all eight vertices, the set of all such triples will be denoted by $\mathcal{P}$.

To analyse the paradoxical outcomes and suggest restrictions to overcome them, we will use a profile decomposition technique developed by Saari (1995). From a majority point of view it is clear that two opposite valuations about an issue do cancel out, i.e. have no impact on the majority outcome. This can, however, be extended to any number of valuations by decomposing a profile into subprofiles:

Definition 3 For any profile $\mathbf{p} = (p_1, \ldots, p_{|X|}) \in [0,1]^{|X|}$ with $\sum_k p_k = 1$ a subprofile is a vector $\underline{\mathbf{p}} = (\underline{p}_1, \ldots, \underline{p}_{|X|}) \in [0,1]^{|X|}$ such that $\underline{p}_k \leq p_k$ for all $k \in \{1, \ldots, |X|\}$.

It is obvious that the above decomposition argument for two opposite valuations does hold for any subprofile $\underline{\mathbf{p}}$ of $\mathbf{p}$ for which $x(\underline{\mathbf{p}}) = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. Such a subprofile does not influence the majority outcome based on $\mathbf{p}$ at all.

As an example consider two individuals with the respective valuations v_2 and v_7. They are exact opposites, so from a majority point of view those two valuations cancel out. Hence this implies that in any profile $\mathbf{p}$, for all opposite valuations we can cancel the share of the valuation held by the smaller number of individuals (and correct for the other shares accordingly) and still have the majority outcome unchanged.

Lemma 1 *Let* $\mathbf{p}$ *and* p' *be two profiles such that, for* $i \in \{1, \ldots, 8\}$,

$$p'_i = \frac{max\{p_i - p_{9-i}, 0\}}{\sum_{k=1}^{4} |p_k - p_{9-k}|}.$$

Then $x^j(\mathbf{p}) \geq \frac{1}{2} \Leftrightarrow x^j(\mathrm{p}') > \frac{1}{2}$.

[2] Equivalently we could say that they are each of Hamming distance 1 from their common neighbor.

Table 3 Triples of neighbors

Valuation	Triple	Valuation	Triple
v_1	(v_2, v_3, v_4)	v_5	(v_2, v_3, v_8)
v_2	(v_1, v_5, v_6)	v_6	(v_2, v_4, v_8)
v_3	(v_1, v_5, v_7)	v_7	(v_3, v_4, v_8)
v_4	(v_1, v_6, v_7)	v_8	(v_5, v_6, v_7)

Proof The average support for each of the three issues can be stated as follows:

$$x^1(\mathbf{p}) = p_2 + p_5 + p_6 + p_8$$
$$x^2(\mathbf{p}) = p_3 + p_5 + p_7 + p_8$$
$$x^3(\mathbf{p}) = p_4 + p_6 + p_7 + p_8$$

Now, let $x^j(\mathbf{p}) = a$ and for some i, $|p_i - p_{9-i}| = t$, and assume w.l.o.g. that $1 > a \geq \frac{1}{2}$ and $0 < t < a$. For $\frac{a}{1} \geq \frac{1}{2}$ we also get $\frac{a-t}{1-2t} \geq \frac{1}{2}$. To see this suppose this is not the case, i.e. $\frac{a-t}{1-2t} < \frac{1}{2}$. It follows that $2a - 2t < 1 - 2t$. For $a \geq \frac{1}{2}$ this is false and therefore $\frac{a-t}{1-2t} \geq \frac{1}{2}$ is true. Repeat this for all $i \in \{1, \ldots, 4\}$. For necessity just reverse the above arguments. □

The lemma shows that in p′ at most 4 entries can be positive. As already previously mentioned, we can reduce a profile by any subprofile that does not change the majority outcome. Consider a subprofile $\underline{\mathbf{p}}$ with positive shares only for the valuations $(0,0,0), (1,1,0), (1,0,1)$ and $(0,1,1)$, namely $\underline{\mathbf{p}} = (\frac{1}{8}, 0, 0, 0, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, 0)$. On each issue there is the same number of individuals in favor of it and against it, i.e. $x(\underline{\mathbf{p}}) = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. Hence, the elimination of such a subprofile does not change the majority outcome of the original profile and eventually increases the number of zero entries in the profile. The only two sets of valuations useable for such a reduction are $\{v_1, v_5, v_6, v_7\}$ and $\{v_2, v_3, v_4, v_8\}$.

Both, the pairwise reduction as well as the reduction using 4 valuations, lead to a reduced profile, the majority outcome of which is identical to the majority outcome of the original profile.

Now we can use the above concepts for a result in a three-dimensional framework, namely that the Anscombe paradox manifests itself in a particularly strong form:

Proposition 3 For $|J| = 3$ the Anscombe paradox will always show up in its *strong form*, i.e. a majority of the voters has a lower Hamming distance to the valuation which is the exact opposite of the majority outcome than to the majority outcome itself.

Proof Assume, w.l.o.g., that we want the majority outcome to be ${}^M v(\mathbf{p}) = (0,0,0)$. As $|J| = 3$, each voter k among a majority of the voters needs to have $d_H(v_k, {}^M v(\mathbf{p})) \geq 2$. Starting with ${}^M v(\mathbf{p}) = v_1$ this leads to the following conditions needed to be satisfied for the Anscombe paradox to occur, where the first three conditions guarantee the majority outcome to be v_1 and the fourth condition ensures that a majority of voters is of a Hamming distance of at least 2 from the majority outcome:

- $x^1(\mathbf{p}) = p_2 + p_5 + p_6 + p_8 < \frac{1}{2}$
- $x^2(\mathbf{p}) = p_3 + p_5 + p_7 + p_8 < \frac{1}{2}$
- $x^3(\mathbf{p}) = p_4 + p_6 + p_7 + p_8 < \frac{1}{2}$
- $p_5 + p_6 + p_7 + p_8 > \frac{1}{2}$

Based on our previous decomposition argument (especially Lemma 1), no identical change in p_8 and p_1 would change the truth of any of the above inequalities. But this is also true for any other pair of opposite valuations. Hence we can directly look at the reduced profile p' with at most 4 entries. For ${}^{M}v^{j}(textp') = 0$ it is not possible that more than half of the shares are located on one plane of the cube, i.e. $p'_8 + p'_r + p'_s < \frac{1}{2}$ for all $r, s \in \{5, 6, 7\}$. But this implies that $p'_r > 0$ for all $r \in \{5, 6, 7\}$ and hence $p'_1 > 0$ (and therefore $p'_8 = 0$ to enable ${}^{M}v(\mathrm{p}') = \mathrm{v}_1$). Now for any $k \in \{5, 6, 7\}$, v_k is not closer to a majority of the voters' valuation than to ${}^{M}v(\mathbf{p})$, as $p_k < \frac{1}{2}$, and this would be the only voters with smaller distance. For any $k \in \{2, 3, 4\}$, v_k is not closer to a majority of the voters' valuation than to ${}^{M}v(\mathbf{p})$, as $p'_r + p'_s < \frac{1}{2}$ for all $r, s \in \{5, 6, 7\}$ and only two valuations out of $\{v_5, v_6, v_7\}$ are closer to v_k than to ${}^{M}v(\mathbf{p})$. $\square$

4 Judgment Aggregation and the Logical Inconsistency of the Majority Outcome

In judgment aggregation, the issues in the agenda are logically interconnected propositions and thus not all valuations are admissible, i.e. logically consistent. Given the binary structure of the problem, we see that the tools of the geometric approach can be used to analyse paradoxes of judgment aggregation.[3] The discursive dilemma with the agenda $\{p, q, p \wedge q\}$ and the associated set of admissible valuations $X = \{(0,0,0), (1,0,0), (0,1,0), (1,1,1)\}$ can again be analysed in our three-dimensional hypercube, in which the four admissible vertices determine the representation polytope as seen in Fig. 2.

Given the set of admissible valuations X, consider the profile $\mathbf{p} = (0, \frac{1}{3}, \frac{1}{3}, 0, 0, 0, 0, \frac{1}{3})$, i.e. no voter has valuation $(0,0,0)$, one third of the voters has valuation $(1,0,0)$, and so on. As this maps into the point $x(\mathbf{p}) = (\frac{2}{3}, \frac{2}{3}, \frac{1}{3})$—a point whose closest vertex is $(1,1,0)$—the representation polytope obviously passes through the majority subcube of an inadmissible valuation, i.e. the set of admissible valuations X is not closed under majority voting. In this case X is called majority inconsistent.

[3] See e.g. Saari (2008) for a very brief discussion of the link of his geometric approach to judgment aggregation.

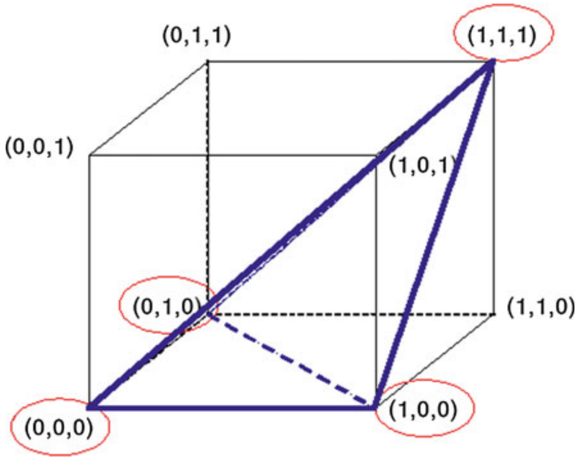


Fig. 2 Representation polytope

That this type of paradox can easily occur with majority voting is seen from the following lemma:

Lemma 2 *Given any vertex* $v \in \{0,1\}^{|J|}$, *there exist 3 vertices* v_a, v_b, v_c *with respective shares* p_a, p_b, p_c *such that for some profile* $\mathbf{p}$ *with* $p_k = 0$ *for all* $k \notin \{a,b,c\}$, $x(\mathbf{p})$ *lies in the v-subcube.*

For $|J| = 3$, *these 3 vertices necessarily need to have v as their common neighbor. Given that, we can now provide a simple result for the majority inconsistency of a set of valuations X, i.e. a necessary condition for X not to be closed under majority voting.*

Proposition 4 *For* $|J| = 3$, *the set of admissible valuations X is majority inconsistent only if for some triple of vertices in X with a common neighbor, this common neighbor is not contained in X.*

In our 3-dimensional setting, we can easily specify all possible triples that could lead to inadmissible majority outcomes. The reduced profile does have an interesting feature in exactly those situations when inadmissible majority outcomes could occur:

Proposition 5 *For* $|J| = 3$, *if for some* $v_i \in \{0,1\}^3$ *with* $p_i \leq p_{9-i}$, *each valuation in the triple of neighbors has a larger share in* $\mathbf{p}$ *than its opposite valuation, then the reduced profile* $\bar{\mathrm{p}}$ *has at most 3 positive entries.*

Proof Let $\mathbf{p} = (p_1, p_2, \ldots, p_8)$ s.t. $\sum_k p_k = 1$. From Lemma 1 we know that

$$p'_i = \frac{max\{p_i - p_{9-i}, 0\}}{\sum_{k=1}^{4} |p_k - p_{9-k}|}.$$

Now in p′ there are at most 4 positive entries. Given that it is not possible that $\left[p'_i > 0 \wedge p'_{9-i} > 0\right]$ for any $i = 1, \ldots, 8$, and that for some v_k each valuation in the

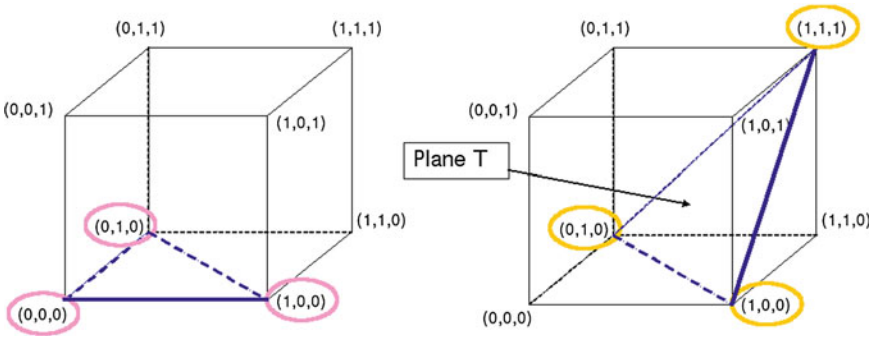


Fig. 3 Planes

triple of neighbors has a larger share than its opposite valuation, this only leaves two possibilities, namely that we have positive shares at most either for all of (p_1, p_5, p_6, p_7) or for all of (p_2, p_3, p_4, p_8). However, in both cases—as was discussed before—further reductions are possible by looking for particular subprofiles. Let—for the above two combinations—$\mathcal{A} = \{i : p_i' > 0\}$ be all valuations for which there is a positive share. Then we can reduce the profile further to profile $\bar{\mathrm{p}}$ such that

$$\bar{p}_i = \frac{max_{j \in \mathcal{A}/\{i\}}\{p_i' - p_j', 0\}}{\sum_{i \in \mathcal{A}} p_i' - min_{j \in \mathcal{A}} p_j'}.$$

Obviously $\bar{\mathrm{p}}$ has at most 3 positive entries. □

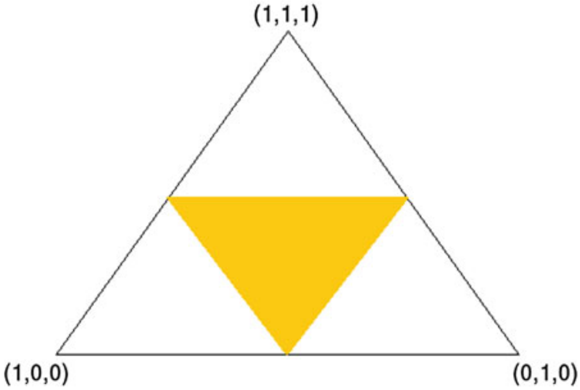
Example 1 Let us consider the following set of admissible valuations $X = \{v_1, v_2, v_3, v_8\}$, i.e. any profile $\mathbf{p} = (p_1, p_2, p_3, 0, 0, 0, 0, p_8)$, where $p_k \geq 0$ for all $k = 1, 2, 3, 8$ and $\sum_k p_k = 1$. As $v_1 = (0, 0, 0)$ and $v_8 = (1, 1, 1)$ are exact opposites, the reduced profile will have a share of 0 for the valuation held by the smaller number of individuals. In the case of $p_1 > p_4$ such a reduced profile will be $\bar{\mathrm{p}} = (\frac{p_1 - p_4}{p_1 + p_2 + p_3 - p_4}, \frac{p_2}{p_1 + p_2 + p_3 - p_4}, \frac{p_3}{p_1 + p_2 + p_3 - p_4}, 0, 0, 0, 0, 0)$, in the case of $p_1 \leq p_4$ we can create the reduced profile accordingly. Hence the reduced profile maps into one of the following two planes shown in Fig. 3, namely either into the one determined by the vertices v_1, v_2 and v_3 or the one determined by the vertices v_2, v_3 and v_8.

Problems may arise if the reduced profile has positive shares only for 3 valuations that constitute a triple in $\mathcal{P}$. For the above example this would be the triple (v_2, v_3, v_8) on the right side of Fig. 3. Now, in Fig. 4 the intersection of this plane with the $(1, 1, 0)$ majority subcube is indicated by the shaded triangle. Only if the reduced profile maps into this triangle do inadmissible majority outcomes arise.

Now, for the 3-dimensional framework we can state the following result:

Proposition 6 *A set of admissible valuations* $X \subseteq \{0, 1\}^3$ *is majority inconsistent if and only if for some reduced profile* $\bar{\mathrm{p}}$ *the following conditions are met*:

Fig. 4 Plane T



- $\bar{p}$ has 3 positive entries
- the 3 valuations with positive shares form a triple $(v_a, v_b, v_c) \in \mathcal{P}$ whose common neighbor is not in X
- the following condition holds for all $v_k \in \{v_a, v_b, v_c\}$ with corresponding shares $\bar{p}_k \in \{\bar{p}_a, \bar{p}_b, \bar{p}_c\}$:

$$\frac{\bar{p}_k}{\bar{p}_a + \bar{p}_b + \bar{p}_c} \leq \frac{1}{2}$$

Proof The sufficiency part is obvious from Fig. 4. For necessity, it is clear that with less than 3 positive entries in $\bar{p}$ no inadmissible outcome can occur. Moreover, any triple not in $\mathcal{P}$ is closed under majority rule. In the case of 4 positive entries in $\bar{p}$ problems only arise in case a triple in $\mathcal{P}$ has a positive share where the common neighbor is not contained in X. But then the fourth positive entry must be one such that the resulting profile can still be further reduced and hence this contradicts the assumption that $\bar{p}$ was the reduced profile already. Now, the only further option is a triple in $\mathcal{P}$ with a common neighbor not in X. In this situation inconsistency occurs exactly in the cut with the respective majority subcube (see Fig. 4) whose points are specified by the conditions above. $\square$

One interesting feature of this result is that the complementary set of profiles actually determines the domain that is closed under majority rule. As those restrictions are based on the space of profiles, this approach is more general than restrictions on the space of valuations which is usually used in the classical literature on domain restrictions. For example List (2005) introduces the unidimensional alignment domain which has a certain resemblance to Black's single peakedness condition in social choice theory. It requires individuals to be ordered from left to right such that on each proposition there occurs only one switch from believing it to not believing it (or vice versa). For $|J| = 3$ a unidimensional alignment domain would not satisfy one of the above conditions for inadmissible majority outcomes.[4]

In addition, this geometric framework also opens a simple way to analyse various other paradoxical situations, e.g. strong support for one particular issue combined with an inadmissible majority outcome, as given in the following proposition:

Proposition 7 *There exist profiles such that there is almost unanimous agreement on one issue and still an inadmissible majority outcome is obtained.*

Proof Looking at Fig. 4 one observes, that points close to the edge connecting the vertices $(1,0,0)$ and $(1,1,1)$ have almost unanimous agreement on issue 1. However, at the midpoint of this edge, the shaded triangle comes arbitrarily close to the edge. Hence, there exist profiles which lie in the shaded triangle but imply almost unanimous agreement on one issue. The same argument applies to points close to the edge connecting the vertices $(0,1,0)$ and $(1,1,1)$.□

5 Likelihood of Inadmissible Majority Outcomes

The geometric framework can also be used to analyze the likelihood of inadmissible majority outcomes in case $|X| \leq 4$. The approach is based on the fact that only 4 vertices are admissible individual valuations, and hence any point $x(\mathbf{p})$ in the representation polytope is determined by a unique profile $\mathbf{p}$. Consider again the situation $X = \{(0,0,0),(1,0,0),(0,1,0),(1,1,1)\}\}$. Then for any vector of shares of individual valuations $\mathbf{p} = (p_1, p_2, p_3, 0, 0, 0, 0, p_8)$ we get the following average support on each issue: $x^1(\mathbf{p}) = p_2 + p_8$, $x^2(\mathbf{p}) = p_3 + p_8$, $x^3(\mathbf{p}) = p_8$, $1 = p_1 + p_2 + p_3 + p_8$. As those are 4 equations with 4 unknowns there exists a unique solution. Thus, assuming every profile being equally likely—i.e. taking an impartial anonymous culture[5]—the volume of certain subspaces now indicates the likelihood of occurrence of certain outcomes. Consider first the volume of the representation polytope V_R: $V_R = \frac{1}{2} \cdot 1 \cdot \frac{1}{3} = \frac{1}{6}$. On the other hand, points leading to inadmissible majority outcomes are located in the tetraeder determined by the points $\left[(\frac{1}{2},\frac{1}{2},0),(1,\frac{1}{2},\frac{1}{2}),(\frac{1}{2},1,\frac{1}{2}),(\frac{1}{2},\frac{1}{2},\frac{1}{2})\right]$. The volume of this tetraeder, V_T, is $V_T = \frac{1}{48}$ (see Fig. 5).

Now, the volume of the tetraeder relative to the volume of the whole representation polytope is $\frac{V_T}{V_R} = \frac{1}{8}$ and hence we can say that the probability of a majority outcome being inadmissible is 12.5 %. This provides another—geometric—approach to compute the probability of paradoxical situations under impartial anonymous culture, leading to the same results as a previous approach by List (2005).

[4] For a more elaborated discussion on majority voting on restricted domains see also Dietrich and List (2010).

[5] See Gehrlein (2006) for a general discussion of the impartial anonymous culture.

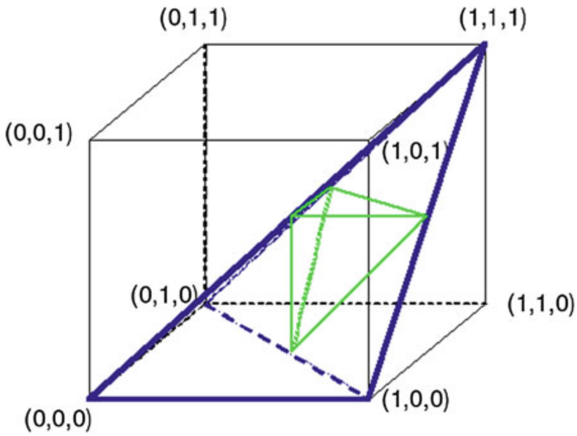


Fig. 5 Distances

Of course, different domains allow for different probabilities. For example consider the agenda $\{p, q, p \leftrightarrow q\}$ with $X = \{(1,0,0), (0,1,0), (0,0,1), (1,1,1)\}$. Then, for any point $x(\mathbf{p}) = (x^1, x^2, x^3)$ in the representation polytope we get $x^1(\mathbf{p}) = p_2 + p_8, x^2(\mathbf{p}) = p_3 + p_8, x^3(\mathbf{p}) = p_4 + p_8$ and $p_2 + p_3 + p_4 + p_8 = 1$. Again, every profile maps into a unique point in the representation polytope. Making the same volume calculations as before, we get—under the impartial anonymous culture—a probability of inadmissible majority outcomes of 25 %.

6 Codomain Restrictions and Distance-Based Aggregation

We saw that using majority rule may lead to inadmissible majority outcomes. Restrictions on the space of profiles are one possibility to overcome such problems. An alternative way to guarantee admissible majority outcomes is to restrict the set of collective outcomes to admissible valuations.[6] One way to work with such codomain restrictions is by using distance-based aggregation rules. Meskanen and Nurmi (2006) have provided an extensive analysis of distance-based aggregation rules in the Arrovian framework. In analogy to a well-known procedure in social choice theory (Kemeny 1959), Pigozzi (2006) introduced such an approach to judgment aggregation. In principle a distance-based aggregation rule determines the collective valuation as the admissible valuation that minimizes the sum of Hamming distances to the individual valuations. Formally, this can be stated as follows:

$$f(\mathbf{p}) = \arg\min_{v \in X} \sum_{k=1}^{|X|} p_k d_H(v, v_k)$$

[6] In social choice theory, aggregation rules based on such restrictions are often called Condorcet extensions.

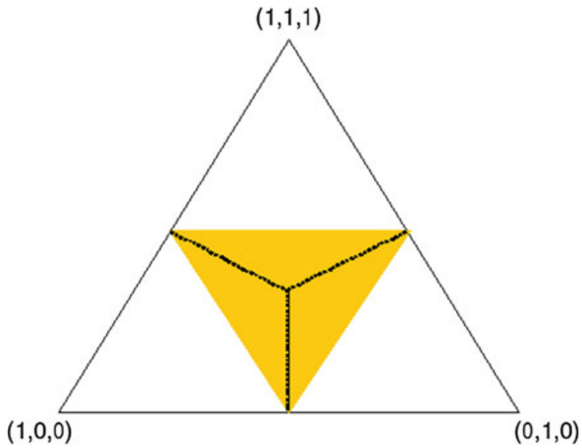


Fig. 6 Distances

Given our geometric approach, there is a simple geometric explanation of this distance-based aggregation rule. As could be seen in Fig. 4, all toublesome profiles lead to a point of average support in the shaded triangle. However, one option is to divide the triangle into three sub-triangles as in Fig. 6.

The intersection point of the three lines is exactly the barycenter point of the triangle, $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. Those lines divide the shaded triangle into three equally sized sub-triangles, points in each sub-triangle are characterized by being of smallest Euclidean distance to the same vertex of the large triangle. For example points in the south-western sub-triangle will be closest to the $(1, 0, 0)$ vertex. As has been shown by Merlin and Saari (2000), the same will be obtained if—for any point in the shaded triangle—one switches the majority valuation on the issue j which is closest to the 50–50 threshold, i.e. for which $x^j(\mathbf{p})$ is closest to $\frac{1}{2}$. This will be illustrated in the following example:

Example 2 Let $X = \{(0,0,0), (1,0,0), (0,1,0), (1,1,1)\}$ and $\mathbf{p} = (0.1, 0.35, 0.3, 0, 0, 0, 0, 0.25)$. This leads to $x(\mathbf{p}) = (0.6, 0.55, 0.25)$ and hence an inadmissible majority outcome ${}^{M}v(\mathbf{p}) = (1, 1, 0)$. Looking at Fig. 4 we see that $x(\mathbf{p})$ lies in the south-western sub-triangle. Thus, according to our distance-based aggregation rule, the outcome will be the admissible valuation $(1, 0, 0)$ as $x(\mathbf{p})$ is closest to the $(1, 0, 0)$ vertex. However, this can also be seen as switching the valuation on the proposition j whose average support $x^j(\mathbf{p})$ is closest to $\frac{1}{2}$. In $x(\mathbf{p})$ this is obviously proposition 2.

7 Conclusion

In this chapter we have shown how geometry can be used to analyse paradoxes occuring under majority voting in the general framework of abstract aggregation. In particular—for the three-dimensional case—we have investigated the

Anscombe paradox, various (impossibility) results in judgment aggregation and distance based aggregation rules. In addition we gave generalized domain conditions characterizing these paradoxes and determined the likelihood of such inadmissible majority outcomes.

Most of the stated results do not easily extend to more than three issues because of problems of dimensionality. For example an agenda with three propositions and their conjunction, like $\{p, q, r, p \wedge q \wedge r\}$, leads to eight admissible valuations, i.e. eight vertices out of the 16 vertices in the four-dimensional hypercube. The extensions of our (domain) restrictions and calculations of the likelihood of the occurrence of paradoxes to those higher dimensions are not obvious and need further work.

References

Brams, S. J., Kilgour, M. D., & Sanver, M. R. (2004). A minimax procedure for negotiating multilateral treaties. In M. Wiberg (Ed.), *Reasoned choices: Essays in Honor of Hannu Nurmi*. Turku: Finnish Political Science Association.

Dietrich, F., & List, C. (2010). Majority voting on restricted domains. *Journal of Economic Theory*, *145*(2), 512–543.

Dokow, E., & Holzman, R. (2010). Aggregation of binary evaluations. *Journal of Economic Theory*, *145*(2), 544–561.

Gehrlein, W. V. (2006). *Condorcet's paradox*. Berlin: Springer.

Kemeny, J. (1959). Mathematics without numbers. *Daedalus*, *88*(4), 577–591.

List, C. (2005). The probability of inconsistencies in complex collective decisions. *Social Choice and Welfare*, *24*, 3–32.

List, C. & Puppe, C. (2009). Judgment aggregation. In P. Anand, P. K. Pattanaik & C. Puppe (Eds.), *The handbook of rational and social choice: An overview of new foundations and applications*. Oxford: Oxford University Press.

Merlin, V., & Saari, D. G. (2000). A geometric examination of Kemeny's rule. *Social Choice and Welfare*, *17*(3), 403–438.

Meskanen, T., & Nurmi, H. (2006). Distance from consensus: A theme and variations. In B. Simeone & F. Pukelsheim (Eds.), *Mathematics and democracy*. Berlin: Springer.

Nurmi, H. (1987). *Comparing voting systems*. Dordrecht: D. Reidel.

Nurmi, H. (1997). Compound majority paradoxes and proportional representation. *European Journal of Political Economy*, *13*(3), 443–454.

Nurmi, H. (1999). *Voting paradoxes and how to deal with them*. Berlin: Springer.

Nurmi, H. (2004). A comparison of some distance-based choice rules in ranking environments. *Theory and Decision*, *57*(1), 5–24.

Pigozzi, G. (2006). Belief merging and the discursive dilemma: An argument-based account to paradoxes of judgment aggregation. *Synthese*, *152*(2), 285–298.

Rubinstein, A., & Fishburn, P. C. (1986). Algebraic aggregation theory. *Journal of Economic Theory*, *38*(1), 63–77.

Saari, D. G. (1995). *Basic geometry of voting*. Berlin: Springer.

Saari, D. G. (2008). *Disposing dictators, demystifying voting paradoxes: Social choice analysis*. New York: Cambridge University Press.

Wilson, R. (1975). On the theory of aggregation. *Journal of Economic Theory*, *10*(1), 89–99.

Necessary and Sufficient Conditions to Make the Numbers Count

Marlies Ahlert and Hartmut Kliemt

1 Introduction

The discussion of whether or not the numbers should count in ethical rankings of states of affairs has been going on for quite a while. It was preceded by related disputes like those over the 'trolley problem' (Thomson 1976), the survival lottery (Harris 1975) or more generally speaking utilitarianism's lack of respect for the separateness of persons [an old concern voiced forcefully in particular by Rawls (1971), see also Kliemt (1998)]. But Taurek's (1977) problem of how to allocate an insufficient supply of a drug when six individuals are doomed if they do not get access to a sufficient quantity of that drug is, arguably, the most instructive one (and not burdened with standard action and omission problems).

In Taurek's example David needs all of the drug to survive while five other persons could each survive on one-fifth of the supply. If in the name of substantive equality we should give nothing or exactly one-sixth to each of the individuals we would be letting them all die. Throwing a fair coin would give each an equal survival chance of one half—with an expected value of the number of survivors of three. Allocating the full supply to David would rescue him for sure. Allocating the full supply in quantities of one-fifth would rescue the five with certainty while ringing the bell on David.

Subsequently we do not intend to argue for or against any specific answer to Taurek's allocation problem. We suggest to look at the problem in a more indirect

This chapter has been published in Homo Oeconomicus 26 (2), 2009.

M. Ahlert (✉)
Department of Law and Economics, Martin-Luther-University Halle-Wittenberg, Halle, Germany
e-mail: marlies.ahlert@wiwi.uni-halle.de

H. Kliemt
Frankfurt School of Finance and Management, Frankfurt, Germany
e-mail: h.kliemt@frankfurt-school.de

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*, DOI: 10.1007/978-3-642-35929-3_34, © Springer-Verlag Berlin Heidelberg 2013

way by shifting the focus towards axioms that characterize the numbers' criterion. When ranking sets of individuals who are rescued or destined to perish the numbers' criterion is implied by three seemingly innocuous axioms. Anybody who does not agree that counting the numbers is the ethically correct way to rank sets of individuals (rescued or doomed) must reject at least one of the three axioms. Therefore discussing the normative appeal of each of the axioms may help us on in finding out what exactly is at stake in this context.

2 Axiomatic Framework[1]

2.1 Preliminary Definitions

We consider an arbitrary but specific finite set X of individuals (human persons) who all will die unless by some external intervention a subset of these individuals be saved. Relevant decision problems will only emerge if X contains at least two persons. Individuals are indicated by small letters eg. x, y, z. They are members (elements) of the finite set X of all individuals under consideration in the context at hand. Subsets of X are denoted by $\{x\}$ or $\{x, y\}$, $\{x, y, z\}$…or A, B, C. The set of all subsets of X including the empty set $\emptyset$ is denoted by $\Pi(X)$. We assume that a decision maker allocating a scarce resource to sets of individuals implicitly forms a (binary) moral preference relation on $\Pi(X) \times \Pi(X)$ that is defined by a subset R of $\Pi(X) \times \Pi(X)$.[2] Given any two finite sets A and B of individuals, the decision maker has to be able to say whether or not she deems it morally at least as good to save A as to save B (completeness). If she weakly prefers to save A over saving B this amounts to $(A, B) \in R$. Instead of $(A, B) \in R$ we often write in the conventional manner $A\ R\ B$. We assume that R is an ordering.[3]

Also in the conventional manner we define strict preference for $A, B \in \Pi(X)$ by $A\ P\ B$: $(A\ R\ B\ \&\ \neg\ B\ R\ A)$ and indifference by $A\ I\ B$: $(A\ R\ B\ \&\ B\ R\ A)$. We note that $A\ R\ B$ actually could be read as 'rescuing A and dooming $X \backslash A$' is weakly preferred to 'rescuing B and dooming $A \backslash B$' since the situation is specified such that an individual is rescued if and only if the individual is included in the chosen set. In the following for a finite set A, $\#A$ denotes the cardinality of A.

1 In this section we transfer the axioms presented by Pattanaik and Xu(1990) to the framework used in the discussion on 'Do the numbers count?'

2 In fact this may not be much of an 'assumption' as long as we have not required anything about the comparison. It is practically implied by any ranking activity.

3 That R forms an ordering is, of course, a substantial assumption. A binary relation R on $\Pi(X) \times \Pi(X)$ is an ordering, if and only if for all $A \in \Pi(X)$: $A\ R\ A$ holds (reflexivity of the ranking R) *and* for all $A, B \in \Pi(X)$: $A\ R\ B$ or $B\ R\ A$ holds (completeness of the ranking R) *and* for all $A, B, C \in \Pi(X)$: $A\ R\ B$ and $B\ R\ C \Longrightarrow A\ R\ C$ holds (transitivity of the ranking R). The proof of the central theorem, however, requires only transitivity. The other properties of an ordering are then implied by the axioms.

2.2 Axioms for Ranking Sets of Individuals

Axiom 1 Indifference between Singletons

For all individuals x, $y \in X$, $\{x\}I\{y\}$.

According to this axiom, for all x, $y \in X$ the decision maker is indifferent between any two situations where she can save exactly one person, either person x or person y. Axiom 1 seems to be an almost unavoidable implication of the universalization requirement in passing moral judgement that is accepted in almost all ethical theory.[4]

Axiom 2 Simple Set Expansion Monotonicity[5]

For all individuals $x \in X$, $\{x\}\ P\ \emptyset$.

This axiom requires that to save one person's life is better than saving no life.[6] It may be seen as expressing minimum beneficence.

Axiom 3 Independence For all A, $B \in \Pi(X)$ and for all $x \in X\backslash\ (A \cup B)$,

$$A\,R\,B \Leftrightarrow A \cup \{x\} R\, B \cup \{x\}.$$

Assume that the decision maker has ranked two sets A and B of individuals, such that she (weakly) prefers to save the lives of the individuals in A to saving the lives of those in B. Now another person x that does not belong to A and not to B joins both groups, i.e. x is saved together with the individuals in A and x is also saved if the individuals in B are saved. Then the ranking between A and B is always identical to the ranking of $A \cup \{x\}$ and $B \cup \{x\}$.Weak preference between A and B remains weak preference between $A \cup \{x\}$ and $B \cup \{x\}$, strict preference remains strict preference and indifference remains indifference.

One can also equivalently read the axiom 'backwards': Assume that in two sets C and $D \in \Pi\ (X)$ such that $C\ R\ D$, there exists the same individual $x \in C \cap D$. Now assume that x cannot be saved any more (e.g. x has died), then this should not change the ranking between the remaining sets $C\backslash\{x\}$ and $D\backslash\{x\}$.

Axiom 3 may be seen as more demanding than the two preceding axioms since certain types of holistic interdependence are ruled out by it. In particular, the view that moving from C to $C\backslash\{x\}$ may harm a group C more than removing x from D such that $D\backslash\{x\}$ emerges with $D\backslash\{x\}\ P\ C\backslash\{x\}$ seems plausible. However, in view of the uniqueness of personality that we ascribe to individuals in *general* it is at least doubtful whether such 'holistic' effects should count for much. Moreover, since

[4] Even if we are ethical non-cognitivist we may still use this axiom as a constitutive characteristic of moral discourse. Generalization in ethics is, of course, classically discussed in Singer (1971), with respect to utilitarianism in Hoerster (1971/1977), while its relation to the very concept of morals is analyzed in Singer (1973).

[5] It can be easily seen that the first axiom along with the premise that there is one individual whom it is better to rescue than not and transitivity implies axiom 2.

[6] Pattanaik and Xu (1990) in ranking sets of objects (rather than sets of human individuals) assumedfor all x, $y \in X$, $x \neq y$, $\{x, y\}\ P\ \{y\}$.

holistic effects should be seen as exceptions rather than the rule, the axiom would still be acceptable for all contexts in which such exceptions do not apply.

In the next step of our argument we will show that the preceding axioms are sufficient and necessary to make the numbers count. We show that the three axioms characterize a ranking of sets of individuals in $\Pi(X)$ that is the same as the one generated by the simple method to count the numbers.

3 Axiomatic Characterization of Counting the Numbers

Definition 1 For all $A, B \in \Pi(X)$, $A\ R_{\#}\ B$: $\Longleftrightarrow \#A \geq \#B$.

Pattanaik and Xu (1990) prove a characterization theorem for the cardinality based ranking of non-empty sets of objects. In their version of the theorem only transitivity of the ranking has to be required, while in their proof reflexivity and completeness are implied by the other assumptions (cf. Barberà, Bossertand Pattanaik 2003). We transfer the theorem to sets of individuals and in that process very slightly generalize the theorem and the proof to Include empty sets as well.

Theorem 1 Let *R be a transitive binary relation on* $\Pi(X) \times \Pi(X)$. *R fulfils Indifference between Singletons, Simple Set Expansion Monotonicity and Independence if and only if* $R = R_{\#}$.

Proof 1. It is obvious that a ranking of sets based on counting and comparing the numbers of individuals in sets is transitive and fulfils the three axioms.

2. Let R be any transitive binary relation on $\Pi(X) \times \Pi(X)$ (recall that the corresponding strict relation is denoted by P and that of indifference by I) satisfying the three axioms. We have to show that for all $A, B \in \Pi(X)$

(2.1) $\#A = \#B \Rightarrow A\ I\ B$ and

(2.2) $\#A > \#B \Rightarrow A\ P\ B$.

(2.1) is proven by induction over the cardinality n of sets.[7]

Let x be any person then $\{x\}\ I\ \{x\}$ by Indifference between Singletons, implying $\{x\}\ R\ \{x\}$ by definition of I. Applying Independence means $\{x\}\ R\ \{x\} \Rightarrow \emptyset\ R\ \emptyset$, implying $\emptyset\ I\ \emptyset$. Thus we have stated the starting point of the induction for $n = 0$ and $n = 1$. To prove the induction step from n to n +1, assume that (2.1) is true for all sets $A, B \in \Pi(X)$ such that $\#A = \#B = n$ with $0 \leq n < \#X$. Consider $A, B \in \Pi(X)$ such that $\#A = \#B = n + 1$ with $0 \leq n < \#X$. Let $C \subset A$ be a subset of A such that $\#C = n$. Then there exists some $x \in A$ such that $A \backslash C = \{x\}$.Obviously $x \in B$ or $x \notin B$.

[7] The start of the induction in the proof by Pattanaik and Xu is $n = 1$.If $\#A = \#B\ A\ I\ B$ is implied by Indifference between Singletons. In our case we have to deal with $\#A = 0$, too.

Case (1) $x \in B$.

In this case define $D = B \backslash \{x\}$. Since $\#C = \#D = n$ we can apply the induction assumption for n to conclude that $C\ I\ D$ holds good. Since x is not contained in $C \cap D$, independence implies

$$A = C \cup \{x\} I\, D \cup \{x\} \ = B.$$

Case (2) $x \notin B$.

Since $x \in A$ and p; $x \notin B$ and $\#A = \#B$, $B \backslash A \neq \emptyset$ has to hold and there exists a person $y \in B \backslash A$. Define $E = B \backslash \{y\}$. $\#C = \#E = n$ and by induction assumption it follows that $C\ I\ E$ holds. Independence implies $A = C \cup \{x\}\ I\ E \cup \{x\}$. Since $\#E > 0$ there exists an element $z \in E$. $\#[(E \cup \{x\}) \backslash \{z\}] = n = \#(B \backslash \{z\})$. The induction assumption implies $(E \cup \{x\}) \backslash \{z\}\ I\ B \backslash \{z\}$. Independence leads to $E \cup \{x\}$ $I\ B$. We have $A\ I\ E \cup \{x\}$ and $E \cup \{x\}\ I\ B$ and by transitivity $A\ I\ B$.

In order to prove (2.2) we choose any two sets $A, B \in \Pi(X)$ such that $\#A > \#B$. We can choose some subset $F \subset A$ such that $\#F = \#B$. (2.1) implies $F\ I\ B$. We define $G = A \backslash F$ which is a nonempty set.

Let g be an element in G, then Simple Set Expansion Monotonicity implies $\{g\}$ $P\ \emptyset$.Let f be in F then Independence implies $\{f, g\}\ P\ \{f\}$. Repeated application of Independence by adding all elements of F stepwise on both sides leads to $\{g\} \cup F$ $P\ F$. Adding now all elements of $G \backslash \{g\}$ stepwise on both sides and applying Independence leads to $A = G \cup F\ P\ F \cup G \backslash \{g\}$. Analogously we receive for any other element h in $G \backslash \{g\}$ by Independence: $F \cup G \backslash \{g\}\ P\ F \cup G \backslash \{g, h\}$ and so on. Transitivity implies $A\ P\ F$ and together with $F\ I\ B$ we receive $A\ P\ B$ which completes the proof. □

As we have shown, the inclusion of the empty set into the framework is possible and the characterization theorem still holds. That in one case we are discussing rankings of sets of human individuals and in one case ranking of sets of objects does not matter in the abstract characterization. Therefore we have now one answer to the question of when the numbers should count: i.e. whenever transitivity of the ranking and axioms 1–3 are accepted. The axioms are also independent as we will show next.

4 Independence of Axioms

This section illustrates the independence and the workings of the three axioms in bringing about the result. If we choose a subset of the axioms by omitting one of them there are possibilities to rank sets of individuals other than by counting numbers.

First we look for a transitive binary relation R_{queen} on $\Pi(X) \times \Pi(X)$ that fulfils Simple Set Expansion Monotonicity and Independence, but not Indifference between Singletons. Let us assume that there is a certain person q (queen or 'David') in the society such that whenever q belongs to a set A this is preferred to any set without q. Sets that either both contain q or both do not contain it are ranked according to numbers of persons. Formally R_{queen} is defined in the following way:

1. For all A, $B \in \Pi(X)$, such that $q \in A$ and $q \in B$ A R_{queen} B: $\Leftrightarrow \#A \geq \#B$.
2. For all A, $B \in \Pi(X)$, such that $q \notin A$ and $q \notin B$ A R_{queen} B: $\Leftrightarrow \#A \geq \#B$.
3. For all A, $B \in \Pi(X)$,such that $q \in A$ and $q \notin B$: A P_{queen} B.

R_{queen} is obviously complete and transitive and thus an ordering. R_{queen} does not fulfil Indifference between Singletons, since for any $x \neq q$, $x \in X$ $\{q\}$ P_{queen} $\{x\}$ holds because of (3).

R_{queen} fulfils Simple Set Expansion Monotonicity. For all $x \in X$, $x \neq q$, $\{x\}$ P_{queen} $\emptyset$ holds because of (2), and $\{q\}$ P_{queen} $\emptyset$ is implied by (3).

R_{queen} fulfils Independence. Let A R_{queen} B be given and $x \in X\backslash$ $(A \cup B)$. If $x = q$, A and B are ranked with respect to numbers (1). $A \cup \{q\}$ and $B \cup \{q\}$ are also ranked with respect to numbers (2). This means that A R_{queen} $B \Leftrightarrow A \cup \{q\}$ R_{queen} B $\cup\{q\}$ holds.

Let $x \neq q$ be given. In cases (1) and (2) of the definition we get the comparison by numbers and hence

A R_{queen} $B \Leftrightarrow A \cup \{x\}$ R_{queen} $B \cup \{x\}$.In case (3) A P_{queen} $B \Leftrightarrow A \cup \{x\}$ P_{queen} B $\cup\{x\}$ holds since $q \in A$ and $q \notin B$.

From this example we can conclude that as soon as we give up the axiom of indifference between Singletons privileging individuals becomes possible. Not only that single persons are treated differently, it is also possible to rank groups according to whether or not a certain person belongs to a group.[8]

Rejecting indifference between Singletons goes against the grain of universalistic ethical theory. If we want to avoid this, what about the other axioms? The next example contains a ranking R_{ind} that fulfils Indifference between Singletons, Independence, but not Simple Set Expansion Monotonicity. Giving up Simple Set Expansion Monotonicity means that there may be cases where saving nobody is at least as good as saving some person. Such a view could be the result of evaluating the tradeoff between equality of treatment or fairness of outcomes for the individuals and efficiency of the decision in favour of equality or fairness (cf. Broome 2002).

For all A, $B \in \Pi(X)$ we define A R_{ind} B, which represents the complete indifference ordering on $\Pi(X)$. If indifference between all sets (including the empty set) applies it is obvious that Indifference between Singletons, Independence, but not Simple Set Expansion Monotonicity are fulfilled.

In addition, we define the inverse ordering $R_{\#inv}$ which compares all A, $B \in \Pi(X)$ by defining A $R_{\#inv}$ $B \Longleftrightarrow \#A \leq \#B$. As a binary relation this ordering is identical to $R_{\#}$, however, the interpretation is inverse.

[8] One could also assign a certain value to each person and rank sets of individuals with respect to the sum of the values assigned to the persons in each set. For the case of ranking sets of opportunities this proposal was modeled by Ahlert (1993).

Proposition 1 *$R_{\#}$, $R_{\#inv}$ and R_{ind} are the only orderings that fulfil Indifference between Singletons and Independence.*

Proof It is easy to show that the three rankings have the desired properties. The inverse direction is shown in two steps.

1. If Simple Set Expansion Monotonicity holds the only candidate is R#.
2. Assume Simple Set Expansion Monotonicity does not hold.

Case (2.1) There is at least one x in X such that $\{x\}\ I\ \emptyset$.We prove the following statement by induction: For all finite sets of cardinality n such that $\#X \geq n > 0$ it holds that they are indifferent to the empty set.

$n = 1$: Indifference between Singletons implies $\{x\}\ I\ \{y\}$ for all $x, y \in X$ therefore by transitivity $\{y\}\ I\ \emptyset$ for all $y \in X$.

Assume the statement holds for $\#X > n \geq 1$ and let a set $A \subseteq X$ with $\#A = n$ +1 be given and $y \in A$, then by induction assumption $A\backslash\{y\}\ I\ \emptyset$. Independence yields $A\backslash\ \{y\} \cup \{y\}\ I\ \emptyset \cup \{y\}$ and this, since $\emptyset \cup \{y\} = \{y\}$ and $\{y\}\ I\ \emptyset$, by transitivity $A\ I\ \emptyset$.Since all sets are indifferent to the empty set transitivity implies that all sets are indifferent. Therefore in case (2.1) the only ranking is R_{ind}.

Case (2.2) For all individuals $x \in X$, $\emptyset\ P\ \{x\}$.This leads to the ordering which is equivalent as a binary relation to $R_{\#}$ but compares inversely. 'Always save the smaller number of individuals'. □

The next counter example is an ordering R_S, that fulfils Indifference between Singletons, Simple Set Expansion Monotonicity, but not Independence. Let us assume that the only important aspect is that lives are saved, but the number of lives does not count. Define

1. for all $A, B \in \Pi(X)$ such that $A \neq \emptyset$ and $B \neq \emptyset$: $A\ I_S\ B$.
2. for all $A, B \in \Pi(X)$ such that $A \neq \emptyset$ and $B = \emptyset$: $A\ P_S\ B$.

R_S is an ordering that fulfils Indifference between Singletons because of (1) and Simple Set Expansion Monotonicity because of (2). R_S does not fulfil Independence, since $\{x\}\ P_S\ \emptyset \Leftrightarrow \{x, y\}\ P_S\ \{y\}$ does not hold for some $y \neq x$ in X which exists, since $\#X > 1$.

5 Conclusion

We have not shown that the numbers should count. We have shown that those who think that the numbers should not count must reject one of three axioms. Rejecting axiom 1 (Indifference between Singletons) seems almost tantamount to giving up universalistic ethical theory. Though ethical particularism may well be the better alternative those who go that route to avoid letting the numbers count should be aware what they are doing. Rejecting axiom 2 (Simple Set Expansion Monotonicity), besides counting, merely allows for complete indifference between all alternatives or always preferring the smaller number. Therefore rejection of

Simple Set Expansion Monotonicity will not lead to much. Rejection of axiom 3 (Independence) as the remaining possibility should attract most attention. Within universalistic ethical approaches (sticking to axiom 1) the relative merits of forming moral judgements in ways conforming with or violating axiom 3 should be discussed if we want to know whether the numbers should count. Ethical universalist who think that the numbers should not count have to be either morally indifferent between all—including zero—numbers of rescued (or doomed) individuals or must reject Independence. The first alternative seems weird while the holistic connotations of the second do not cohere well with the unique value assigned to the individual person which gives rise to Taurek's numbers problem in the first place.

Acknowledgments Hospitality of the economics department of the University of Hamburg is gratefully acknowledged. We are particularly indebted to Mathew Braham, Manfred Holler and Frank Steffen. Comments by Yongsheng Xu on an earlier version helped to clarify some statements. Our general interest in the problem was triggered and inspired by Weyma Lübbe who is Germany's leading expert on the numbers issue. In this chapter we rely strongly on her writings. Particularly useful were Lübbe (2004, 2005).

References

Ahlert, M. (1993). Freedom of choice: a comparison of different rankings of opportunity set. *Social Choice Welfare, 10*, 189–208.

Barberà, S., et al. (2003). Ranking sets of objects, In S. Barberà, et al. (Eds.), *Handbook of Utility Theory*, Kluwer Academic Publishers, Berlin. 2.

Broome, J. (2002) Fairness, goodness, and levelling down, In C. J. L., Murray, et al. (Eds.), *Summary measures of population health. Concepts, ethics, measurement and applications* (pp. 135–137). Geneva.

Harris, J. (1975). The survival lottery. *Philosophy, 50*, 81–87.

Hoerster, N. (1971/1977). *Utilitaristische Ethik und Verallgemeinerung*. Karl Alber.

Kliemt, H. (1998), Rawls's Kritik am Utilitarismus. In H. Otfried (Ed.), *Eine Theorie der Gerechtigkeit*, Akademie Verlag: 97–116.

Lübbe, W.(Ed.) (2004), *Tödliche Entscheidung*. Mentis.

Lübbe, W. (2005), Das Problem der Gleichheit in der 'Numbers' Debatte, in: Rauprich, O. et al. (Eds.), *Gleichheit und Gerechtigkeit in der modernen Medizin,* Mentis, Paderborn: 105–125.

Pattanaik, P., & Yongsheng, Xu. (1990). On ranking opportunity sets in terms of freedom of choice. *Recherches Economiques de Louvain, 56*, 383–390.

Rawls, J. (1971). *A theory of justice*. Oxford: Oxford University Press.

Singer, M. G. (1971). *Generalization in ethics: an essay in the logic of ethics, with the rudiments of a system of moral philosophy*. New York: Russell and Russell.

Singer, P. (1973). The triviality of the debate over 'is: ought' and the definition of 'moral'. *American Philosophical Quarterly, 10*, 51–56.

Taurek, J. (1977). Should the numbers count? *Philosophy and Public Affairs, 6*, 293–316.

Thomson, J. J. (1976). Killing, letting die, and the trolley problem. *The Monist, 59*, 204–217.

Limit Property of a Multi-Choice Value and the Fuzzy Value

Rie Ono-Yoshida

1 Introduction

In ordinary voting games, each player's vote is yes or no. Some researchers have worked on modifications in order to deal with more than two options, and these modifications are divided into two major categories. One is multi-alternative games, defined by Bolger (1986, 1993), which enables players to choose among more than two independent alternatives; e.g., voting in an election in which more than two candidates are running. The other is multi-choice games, defined by Hsiao and Raghavan (1993) and Hsiao (1995), which enable players to choose among more than two participation levels, e.g., voting a yes/no or casting a blank vote. The latter means that the voter is not sufficiently in favor to vote yes, but not sufficiently against to vote no. Both modifications can be discussed not only in the class of voting games but also in broader class of cooperative games. In what follows we will focus on the relationship of multi-choice games and fuzzy games.

In multi-choice games, three values have been proposed as a generalization of the well-known Shapley value (Shapley 1953): those of Hsiao and Raghavan (1993) and Hsiao (1995), of Derks and Peters (1993), and of van den Nouweland et al. (1995). Although the value of Hsiao and Raghavan is derived from a set of axioms, the weight of each participation level must be defined exogenously. To calculate the value of Derks and Peters, we do not need weights, but the value depends on the step number of participation levels, which should be an ordinal number. The value of van den Nouweland et al. is advantageous, because it does not depend on exogenous numbers, and more so, it is capable of an interpretation using permutation like the Shapley value.

This chapter has been published in Homo Oeconomicus 29(3), 2012.

R. Ono-Yoshida (✉)
Faculty of Law and Economics, Chiba University, 1-33 Yayoi-cho Inage-ku,
Chiba 263-8522, Japan
e-mail: onorie@le.chiba-u.ac.jp

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_35, © Springer-Verlag Berlin Heidelberg 2013

Before the concept of multiple-choice gained importance, Aubin (1981, 1993) defined fuzzy games, where each player chooses a participation level in the interval $[0, 1]$. This definition seems closely related to the idea of the multi-choice games. The fuzzy value of Aubin is known as a generalized Shapley value in the fuzzy games.

In this chapter, we discuss two values: the multi-choice value of van den Nouweland et al. and the fuzzy value of Aubin. Our purpose is to show that the multi-choice value is consistent with the fuzzy value when the number of participation levels is sufficiently large. Section 2 describes the basic notation of multi-choice games and fuzzy games, and introduces the multi-choice value of Nouweland et al., and the fuzzy value of Aubin. Section 3 presents a tool, called a piecewise multilinear function, to extend a multi-choice game to a fuzzy game in a natural way. Section 4 contains the main results of this study, which show that the multi-choice value converges to the fuzzy value as the number of participation levels increases. A numerical example illustrates the result. Section 5 concludes.

2 Preliminaries

2.1 Multi-Choice Games

Let us review the multi-choice game defined by van den Nouweland et al. (1995). Let $N = \{1, 2, \ldots, n\}$ be the set of players van den Nouweland (1993) and Each player $i \in N$ chooses a participation level $s_i \in M_i = \{0, 1, \ldots, m_i\}$, where the number of levels $m_i \geq 1$ is an integer. That is, player i has $m_i + 1$ alternatives from which to choose a particular level of participation intensity. The n-tuple

$$s = (s_1, \cdots, s_n) \in \prod\nolimits_{i \in N} M_i$$

is called multi-choice coalition. Function

$$v : \prod\nolimits_{i \in N} M_i \to R \text{with} v(0, \ldots, 0) = 0$$

is called the characteristic function; $v(s)$ is the value that N can gain as a group when the coalition is s. The triple $(N, (M_i)_{i \in N}, v)$ is called a multi-choice game. The set of all multi-choice games with player set N is denoted by MC^N. Note that, if $M_i = \{0, 1\}$ for all $i \in N$, the games in $MC^N$are equivalent to the usual cooperative games.

If N and $(M_i)_{i \in N}$ are well defined, we simply call v a multi-choice game. We assume that every multi-choice game v is monotonic nondecreasing with respect to s. In the following discussion, we also assume that each player has the same set of participation levels so that $M_i = M = \{0, 1, \ldots, m\}$ for all $i \in N$; and the set of all coalitions M^N.

The analogue of unanimity games for multi-choice games is minimal-effort games u_t defined by

$$u_t(s) = \begin{cases} 1, & \text{if } s_{\mathrm{i}} \geq t_i \text{ for all } i \\ 0, & \text{otherwise} \end{cases}$$

where $s, t \in M^N$. We call t_i player i's required level of u_t. This means that the group gets 1 if every player $i \in N$ chooses a level higher than or equal to $t_i \in M$. Every multi-choice game v is described as a linear combination of u_t. This is an extended version of the well-known theorem that every vector is expressed as a linear combination of mutually orthogonal unit vectors.

For arbitrary v, dividend $\Delta_v(s)$ is given *recursively* by

$$\Delta_v(0,\ldots,0) = 0, \text{ and}$$
$$\Delta_v(s) = v(s) - \sum_{r \leq s: r \neq s} \Delta_v(r).$$

This dividend corresponds to the coefficient of the linear combination, that is,

$$v(s) = \sum_{t \in M^N} \Delta_v(t) u_t(s).$$

Note that $\Delta_v(s) < +\infty$, because

$$\Delta_v(s) = v(s) + (n-1)v(s-I) - \sum_{i \in N} v(s - I + e^i),$$

where e^i is the unit n-vector whose ith component equals 1, and $I = (1,\ldots,1)$. We also note that

$$\sum_{s \in M^N} \Delta_v(s) = v(m,\ldots,m).$$

2.2 Generalized Shapley Value

Van den Nouweland et al. (1995) proposed a generalized Shapley value. Let us define an order with m^n elements by bijection $\sigma : N \times (M - \{0\}) \to \{1,\ldots,m^n\}$. An order σ is said to be admissible if it satisfies $\sigma(i,j) < \sigma(i,j+1)$ for all [1] $i \in N$ and $j \in \{1,\ldots,m-1\}$; then there are

$$\frac{(mn)!}{(m!)^n}$$

admissible orders. The set of all admissible orders for a game v is denoted by $\Xi(v)$.

[1] For other solution concepts for multi-choice games such as a core, see van den Nouweland(1993), van den Nouweland et al.(1995), and Branzei et al.(2005).

Given an admissible order σ, let kth coalition be $s^{\sigma,k}$, where

$$s_i^{\sigma,k} = \max\{j \in M | \sigma(i, j) < k\} \cup \{0\}$$

for all $i \in N$, and the marginal contribution of i given j be

$$w_{i,j}^{\sigma} = v(s^{\sigma,\sigma(i,j)}) - v(s^{\sigma,\sigma(i,j)-1})$$

for all $i \in N$ and $j \in \{1, \ldots, m\}$.

Definition 1 (van den Nouweland et al. 1995) Let $v \in MC^N$. The multi-choice value $\varphi(v)$ is the expected marginal contribution of v over all admissible orders, i.e.,

$$\varphi_{i,j}(v) = \frac{(m!)^n}{(mn)!} \sum_{\sigma} w_{i,j}^{\sigma}$$

for all $i \in N$ and $j \in \{1, \cdots, m\}$.

This value is defined by an $m \times n$ matrix. In this chapter, adding a column of figures, let us define the total multi-choice value $\Phi(v) = (\Phi_1, \ldots, \Phi_n)$ as

$$\Phi_i(v) = \sum_{j=1}^{m} \varphi_{i,j}(v).$$

2.3 Fuzzy Games

Aubin (1981, 1993) discussed a fuzzy generalization of the ordinary cooperative games. Let $N = \{1, \ldots, n\}$ be the set of players. Each player $i \in N$ chooses among participation ($s_i = 1$), nonparticipation ($s_i = 0$), and fuzzy participation ($s_i \in (0, 1)$). The n-tuple $s = (s_1, \ldots, s_n) \in [0, 1]^N$ is called the fuzzy coalition. The fuzzy game is the pair (N, v^F), where $v^F : [0, 1]^N \to R$ is continuously differentiable and satisfies $v^F(0, \ldots, 0) = 0$. The set of all fuzzy games with player set N is denoted by FG^N.

Aubin (1981) developed the concept of generalized fuzzy games from an axiomatic approach, similar to Shapley (1953).

Definition 2 (Aubin 1981) Let $v^F \in FG^N$. The fuzzy value is defined by

$$\Theta(v^F) = \int_0^1 \nabla v^F(t, \ldots, t)dt,$$

where $\nabla v^F(\cdot)$ is the gradient of v^F.

In other words, the fuzzy value evaluates the gradient of v^F only on the main diagonal of n-cube $[0, 1]^N$.

3 Fuzzy Games with Piecewise Multilinear Functions

In this section, we present a way to derive a fuzzy game $v^F \in FG^N$ from a multi-choice game $v \in MC^N$. Without loss of generality, renumber the participation levels of a multi-choice game so that the highest level m equals $\frac{m}{m} = 1$, i.e., $M = \{0, \frac{1}{m}, \frac{2}{m}, \ldots, 1\}$. Then, the multi-choice game v is defined only on the grid points in n-cube $[0, 1]^N$. We fill the in-between space of this cube using multi linear functions.

Define $v^F(s) = v(s)$ if s is on a grid point; otherwise, define $v^F(s)$ using a piecewise multilinear function $z : M^N \to [0, 1]^N$. If we divide every edge of a unit n-cube into m equal parts, there would exist a unique small n-cube for which the length of each side equals $\frac{1}{m}$, which includes s. Since we have already defined $v^F(s)$ for each point of this small cube, define $v^F = zv$ by

$$zv(s) = v(x) + m^n \sum_{T \subseteq N} \prod_{j \in T} (s_j - x_j) \prod_{j \notin T} (x_j - s_j + \frac{1}{m}) \left[v\left(x + \frac{e^T}{m}\right) - v(x) \right],$$

where $x_i \in \{0, \frac{1}{m}, \frac{2}{m}, \ldots, \frac{m-1}{m}\}$ is such that $x_i < s_j < x_{i+1/m}$ for all $i \in N$, and e^T is an n-tuple such that

$$e_i^T = \begin{cases} 1, & i \in T \\ 0. & i \notin T \end{cases}$$

We will call $v^F = zv$ a fuzzy game with the piecewise multilinear function extended from a multi-choice game$v$. Figure 1 illustrates the graphical image of the piecewise multilinear function derived from the two-person minimal effort game u_t, if their required level is $t = (\frac{2}{4}, \frac{3}{4})$ as an example.

To clarify the idea of piecewise multilinear functions, see the following three examples.

Example 1 If the number of participation levels is only 1, the multi-choice game v coincides with an ordinary cooperative game. The piecewise multilinear function extended from this game is shown as

$$zv(s) = \sum_{T \subseteq N} v(e^T) \prod_{j \in T} s_j \prod_{j \notin T} (1 - s_j),$$

which is equal to a multilinear extension of an ordinary cooperative game, defined by Owen (1972). Hence, the fuzzy value, which evaluates the gradient of v on the main diagonal of $[0, 1]^N$, coincides with the Shapley value of the cooperative game.

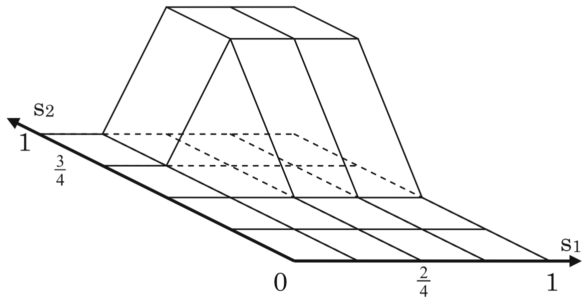


Fig. 1 Piecewise multilinear function derived from u_t, where $t = (\frac{2}{4}, \frac{3}{4})$

Example 2 Let us compute the fuzzy value of the fuzzy game extended from two-person $(m+1)$-choice minimal-effort game u_t, where $t = (t_1, t_2)$. Let us first assume $t_1 > t_2$. There exists $(x_j, x_k) \in M^2$ such that $x_j < t_1 \le x_{j+1/m}$ and $x_k < t_2 \le x_{k+1/m}$ When we evaluate $\nabla zu_t = (\nabla_1 zu_t, \nabla_2 zu_t)$ on the main diagonal of $[0, 1]^2$, gradient $\nabla_1 zu_t$ equals m on $(x_j, x_{j+1/m}] \times (x_j, x_{j+1/m}]$; otherwise, it equals 0. On the other hand, $\nabla_2 zu_t$ equals 0 anywhere on the main diagonal of $[0, 1]^2$. Let us next assume $t_1 = t_2$. Then the elements of gradient $\nabla_1 zu_t$ and $\nabla_2 zu_t$ on $(x_j, x_{j+1/m}] \times (x_j, x_{j+1/m}]$ equal ms_1 and ms_2, respectively. Thus,

$$\Theta(zu_t) = \begin{cases} (1, 0) & t_1 > t_2, \\ (\frac{1}{2}, \frac{1}{2}) & t_1 = t_2, \\ (0, 1) & t_1 < t_2. \end{cases}$$

Example 3 The fuzzy value for n-person multi-choice minimal-effort game u_t is obtained analogously. Let $H(u_t)$ be the set of players who are required for the highest participation level in u_t: i.e.,

$$H(u_t) = \{i \in N | \forall k \in N, t_i \ge t_k\}.$$

Then the fuzzy value for player i is

$$\Theta_i(zu_t) = \begin{cases} \frac{1}{|H|}, & i \in H(u_t) \\ 0. & i \notin H(u_t) \end{cases}$$

Since Θ is a linear operator, the fuzzy value for games zv, extended from general multi-choice games v, is written as a linear combination of the values given in Example 3.

4 Main Theorem

In this section, we discuss the limit property of the multi-choice value. Denote the multi-choice game with $m + 1$ participation levels by v^m, the multi-choice value of the game by $\varphi^m(v^m)$, and the total multi-choice value of the game by $\Phi^m(v^m)$, to clarify the number of participation levels.

Theorem 1 *Let v^m be an $(m + 1)$-choice game and zv^m be the related fuzzy game with the piecewise multilinear function. Then, for every $\varepsilon > 0$ there exists m_ε such that, for all $i \in N$,*

$$\sup_v |\Phi_i^m(v^m) - \Theta_i(zv^m)| < \varepsilon \text{ for all } m > m_\varepsilon.$$

Before we prove this theorem, let us consider two-person $(m + 1)$-choice minimal-effort game u_t, defined in Sect. 2.1, and the related fuzzy game zu_t. For $r, s \in M = \{0, 1/m, \ldots, 1\}$, define $f(r, s)$ as

$$f(r, s) = \frac{(m!)^2}{(2m)!} \cdot \frac{(mr + ms - 1)!}{(mr - 1)!(ms - 1)!} \cdot \frac{(2m - mr - ms)!}{(m - mr)!(m - ms)!}.$$

Note that both mr and ms are integers. Since the total multiple-choice value for the minimal-effort game u_t is written as

$$\Phi_i^m(u_t) = \frac{(m!)^n}{(nm)!} \sum_{s \in L(u_t)} \frac{(\sum_{k \in N} s_k - 1)!}{(\prod_{\substack{k \in N \\ k \neq i}} s_k!) \cdot (s_i - 1)!} \cdot \frac{(mn - \sum_{k \in N} s_k)!}{\prod_{k \in N} (m - s_k)!},$$

where $L(u_t) = \{s \in M^N : s_i = t_i, \text{ and } s_j \geq t_j \quad \text{for all } j \neq i\}$, the total multi-choice value for two-person $(m + 1)$-choice minimal-effort game u_t is

$$\Phi_1^m(u_t) = \sum_{s_2 \geq t_2} f(t_1, s_2) \text{ and } \Phi_2^m(u_t) = \sum_{s_1 \geq t_1} f(s_1, t_2).$$

Lemma 2

$$\sum_{s=0}^{m} f(r, s) = 1$$

Proof Use the identity

$$(1 - x)^{-m-1} = (1 - x)^{-r}(1 - x)^{-m+r-1},$$

and compare the coefficients of x^m of both sides. □

We interpret $f(r, \cdot)$ as a probability density function. Denote a random variable according to this distribution by X. Note that X is a number in $[0, 1]$. We define $F(t) = \Pr(X \geq t)$. Then, $\Phi_1^m(u_t) = F(t_2)$ and $\Phi_2^m(u_t) = F(t_1)$ hold.

The following lemma is the two-person version of Theorem 1.

Lemma 3 *Consider a two-person* $(m+1)$*-choice game* v^m *and the related fuzzy game* zv^m*. For every* $\varepsilon > 0$ *there exists* m_ε *such that, for all* $i \in N$,

$$\sup_v |\Phi_i^m(v^m) - \Theta_i(zv^m)| < \varepsilon \text{ for all } m > m_\varepsilon.$$

Proof We will first prove this assertion for the class of minimal-effort games u_t defined in Sect. 2.1. Since we already know that the fuzzy value of a two-person minimal-effort game is given as in Example 2, we calculate the total multi-choice value for comparison. Let us calculate the total multi-choice value for player 1. For all $\varepsilon > 0$,

$$\Pr(|X - t_2| > \varepsilon) \leq \frac{E(|X - t_2|^2)}{\varepsilon^2}$$

holds from Chebyshev's inequality. The denominator of the right-hand value

$$E(|X - t_2|^2) = \frac{[2(1 - t_2)m + 2t_2 + 1]t}{(m + 1)(m + 2)} \to 0$$

as $m \to 0$, which implies $\Pr(|X - t| < \varepsilon) \to 1$ as $m \to 0$.

Thus we conclude that, for any $\varepsilon > 0$, there exists m_ε such that

$$m > m_\varepsilon \Rightarrow \Pr(X < t_2) = 1 - F(t_2) < \varepsilon,$$

which means $t_1 < t_2$ ($t_1 > t_2$, resp.) implies that $F(t_2)$ converges to 1 (0, respectively), i.e., the total multi-choice value Φ_1^m converges to 0 (1, respectively), as $m \to 0$.

When $t_1 = t_2$, the efficiency and symmetry property of the total multi-choice value imply

$$\Phi_1^m(u_t) = \Phi_2^m(u_t) = \frac{1}{2},$$

for all m. Comparing with the fuzzy value in Example 2, we obtain $\Phi^m(u_t)$ converges to $\Theta(zu_t)$ as $m \to 0$.

Since z, Φ^m, and Θ are linear operators, and since the dividend $\Delta_v(s) < \infty$, the statement holds for any multi-choice game v^m. □

As we mention in the previous section, a multi-choice game with the set of choices $M = \{0, \frac{1}{m}, \ldots, 1\}$ is defined only on the grid points in n-cube $[0, 1]^N$. The total multi-choice value evaluates each player's contribution on all the paths from the origin to $(1, \ldots, 1)$ to consider his/her expected contribution, while the fuzzy value evaluates the contribution only on the main diagonal of n-cube $[0, 1]^N$. The essential part of the proof is that most of the paths from the origin to $(1, \ldots, 1)$

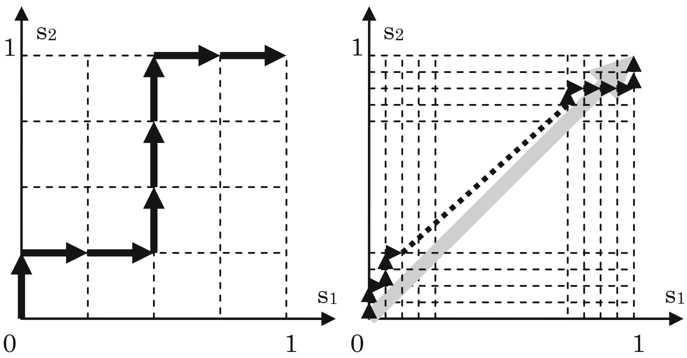


Fig. 2 Evaluation of the total multi-choice value

are close to the main diagonal when m is sufficiently large. Figure 2 provides an overview of these paths for different m.

Denote the number of ways of picking s_1 from s objects by

$$\binom{s}{s_1} = \frac{s!}{s_1!(s - s_1)!}.$$

Let us assume s objects, numbered n boxes ($n \leq s$), and $s_1 + \cdots + s_n = s$. Denote the number of ways of putting s_1 objects into the first box, putting s_2 objects into the second box…, and putting s_n objects into the n-th box by

$$\binom{s}{s_1, \ldots, s_n} = \frac{s!}{s_1! \ldots s_n!}.$$

To prove Theorem 1, it is useful to generalize the function f as

$$f(i : (s_1, \ldots, s_n)) = \frac{(m!)^2}{(nm)!} \cdot \binom{\sum_{k \in N} s_k - 1}{s_1, \ldots, s_i - 1, \ldots, s_n} \binom{mn - \sum_{k \in N} s_k}{m - s_1, \ldots, m - s_n}.$$

Proof of Theorem 1 First, let us consider minimal-effort games and show that the total multi-choice value converges to the fuzzy value as in Lemma 3. Recall that $H(u_t)$ is the set of players who are required for the highest participation level in u_t, and that $\Phi^m(u_t)$ for player $i \notin H(u_t)$ converges to 0. Without loss of generality, assume that player 1 is an element of $H(u_t)$. Then there exists at least one player, say player 2, who is required a higher participation level $t_2 > t_1$. Then,

$$
\begin{aligned}
\Phi_1^m(u_t) &= \sum_{s_j \geq t_j, j \geq 2} f(1 : (s_1, \ldots, s_n)) \\
&= \frac{(m!)^n}{(nm)!} \sum_{s_2 \geq t_2,} \binom{t_1 + s_2 - 1}{s_2} \binom{2m - t_1 - s_2}{m - s_2} \\
&\times \sum_{s_3 \geq t_3, \ldots, s_n \geq t_3} \binom{t_1 + \sum_{j \geq 2} s_j - 1}{t_3, \ldots, t_n} \binom{nm - t_1 - \sum_{j \geq 2} s_j}{m - t_3, \ldots, m - t_n} \\
&\leq \frac{(m!)^n}{(nm)!} \sum_{s_2 \geq t_2,} \binom{t_1 + s_2 - 1}{s_2} \binom{2m - t_1 - s_2}{m - s_2} \\
&\times \sum_{s_3 \geq 0, \ldots, s_n \geq 0} \binom{t_1 + \sum_{j \geq 2} s_j - 1}{t_3, \ldots, t_n} \binom{nm - t_1 - \sum_{j \geq 2} s_j}{m - t_3, \ldots, m - t_n}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{(m!)^n}{(nm)!} \sum_{s_2 \geq t_2,} \binom{t_1 + s_2 - 1}{s_2} \binom{2m - t_1 - s_2}{m - s_2} \times \binom{nm}{2m, m, \ldots, m} \\
&= \frac{(m!)^2}{(2m)!} \sum_{s_2 \geq t_2,} \binom{t_1 + s_2 - 1}{s_2} \binom{2m - t_1 - s_2}{m - s_2} = F(t_2).
\end{aligned}
$$

Since Lemma 3 states that the value $F(t_2)$ for $t_2 > t_1$ converges to 0 as $m \to 0$, the total multi-choice value $\Phi_i^m(u_t)$ for $i \notin H(u_t)$ also converges to 0.

Meanwhile, players in $H(u_t)$ have the same total multi-choice value by symmetry. Using

$$\sum_{i \in N} \Phi_i^m(u_t) = 1,$$

known as the efficiency property, we obtain that the total multi-choice value $\Phi_i^m(u_t)$ for $i \in H(u_t)$ converges to $\frac{1}{h}$. Comparing this with the fuzzy value, calculated as in Example 3, the total multi-choice value converges to the fuzzy value for any minimal-effort game.

Because z, Φ^m, and Θ are linear operators, and the dividend is $\Delta_v(s) < \infty$, the statement of Theorem 1 holds for any multi-choice games v^m. $\square$

Example 4 Let us calculate the total multi-choice value and fuzzy value for two-person $(m + 1)$-choice minimal-effort game u_t, where $t = (\frac{2}{4}, \frac{3}{4})$. The fuzzy multi-

Table 1 Convergence of total multiple-choice value

m	$\Phi_1^m(u_t)$	$\Phi_2^m(u_t)$
4	0.2429	0.7571
8	0.1573	0.8427
20	0.0527	0.9473
100	0.0001	0.9999

choice value is $\Theta(zu_t) = (0, 1)$ as shown in Example 2. When $m = 4$, the graph of the piecewise multilinear function is shown in Fig. 1. As m becomes larger, the total multi-choice value converges to the fuzzy value (0, 1) as shown in Table 1.

5 Conclusion

We discussed the limit property of the multi-choice value proposed by van den Nouweland(1993) and van den Nouweland et al. (1995) and compared this value with the fuzzy value proposed by Aubin (1981, 1993). The sum of multi-choice values over choices, which we called the total multi-choice value, derives from combinatorial interpretation of the well-known Shapley value and depends on the number of levels m. We concluded that the total multiple-choice value is the consistent value in the sense that it connects the Shapley value and the fuzzy value. To obtain this result, we transformed multi-choice game into a fuzzy game, defined the piecewise multilinear function, and demonstrated that the total multiple-choice value converges to the fuzzy value for the extended fuzzy game as m increases.

Finally, I would like to briefly address Bolger's (1993) multi-alternative games and the generalized Shapley value. [1] In multi-choice games, we make admissible orders of both players and alternatives, where all players start by choosing level zero, and then advance step by step from each level to the next one after another, and finally choose level m. This is the key to connect the multi-choice value to the fuzzy value. In contrast, Bolger's generalization of the Shapley value does not assume such orders. However, Ono (2001, 2002) constructs the multilinear extension of the multi-alternative games, which is closely related with a combinatorial interpretation. It focuses on a certain alternative, and assumes the coalition of the players who choose this alternative becomes larger up to the grand coalition. The Bolger value for a player to choose this alternative is the expected contribution of this player over all such coalition-growing processes. The limit properties of this value will be studied in future research.

Acknowledgement The author is grateful to Shigeo Muto, Eiichi Hanzawa, Manfred J. Holler and anonymous referees for their helpful comments.

References

Aubin, J. P. (1981). Cooperative fuzzy games. *Mathematics of Operations Research, 6*, 1–13.

[1] For other solution concepts for multi-choice games such as a core, see van den Nouweland(1993), van den Nouweland et al.(1995), and Branzei et al.(2005).

Aubin, J. P. (1993). *Optima and equilibria*. Berlin: Springer.
Bolger, E. M. (1986). Power indices for multicandidate voting games. *International Journal of Game Theory, 15*, 175–186.
Bolger, E. M. (1993). A value for games with *n* players and *r* alternatives. *International Journal of Game Theory, 22*, 319–334.
Branzei, R., Dimitrov, D., & Tijs, S. (2005). *Models in cooperative game theory: crisp, fuzzy, and multi-choice games*. Berlin: Springer.
Derks, J., & Peters, H. (1993). A Shapley value for games with restricted coalitions. *International Journal of Game Theory, 21*, 351–360.
Hsiao, C. R. (1995). A value for continuously-many-choice cooperative games. *International Journal of Game Theory, 24*, 273–292.
Hsiao, C. R., & Raghavan, T. E. S. (1993). Shapley value for multi-choice cooperative games, I. *Games and Economic Behavior, 5*, 257–267.
Ono, R. (2001). Values for multi alternative games and multilinear extensions. In M. J. Holler & G. Owen (Eds.), *Power indices and coalition formation* (pp. 63–86). Boston: Kluwer.
Ono, R. (2002). Voting systems and power indices: theory and applications, Ph.D. Dissertation, Tokyo Institute of Technology, Tokyo.
Owen, G. (1972). Multilinear extensions of games. *Management Science, 18*, 64–79.
Shapley, L.S. (1953).A value for n-person games.*Annals of Mathematical Studies*,*28*, 307–317.
van den Nouweland, A. (1993). Games and graphs in economic situations, Ph.D. Dissertation, Tilburg University, Tilburg.
van den Nouweland, A., Tijs, S., Potters, J., & Zarzuero, J. (1995). Cores and related solution concepts for multi-choice games. *ZOR-Mathematical Methods of Operations Research, 41*, 289–311.

Pure Bargaining Problems and the Shapley Rule

Francesc Carreras and Guillermo Owen

1 Introduction

The proportional rule has a long tradition in collective problems where some kind of monetary utility (costs, profits, savings...) is to be shared among the agents. However, while its (apparent) simplicity might seem a reason for applying it in pure bargaining affairs, where only the whole and the individual utilities matter, its behavior is, in fact, questionable. A consistent alternative, the *Shapley rule*, will be suggested as a much better solution for this kind of problems (we do not use here the term "consistency" in the specific sense introduced in Hart and Mas-Colell (1989) with regard to a "reduced game" notion). Utilities will be assumed to be completely transferable, so that the class of problems considered here differs from the class of "bargaining problems" more commonly analyzed in the literature.

The organization of the chapter is as follows. In Sect. 2, the notion of *pure bargaining problem* (*PBP*, for short) is provided and the notion of *sharing rule* is stated and exemplified with the classical *proportional rule* and the *equal surplus sharing rule*. In Sect. 3 we attach to any PBP a quasi-additive game (*closure*), thus reducing any PBP to a cooperative game. By using this idea, in Sect. 4 we introduce the *Shapley rule* for PBPs, compare it with the proportional rule, and characterize those PBPs for which the Shapley rule and the proportional rule coincide. Section 5

This chapter has been published in *Homo Oeconomicus* 28(3), 2011.

F. Carreras
Department of Applied Mathematics II and Industrial and Aeronautical Engineering School of Terrassa, Technical University of Catalonia, Catalonia, Spain

F. Carreras (✉)
ETSEIAT, P.O. Box 577, 08220 Terrassa, Spain
e-mail: francesc.carreras@upc.edu

G. Owen
Department of Mathematics, Naval Postgraduate School, Monterey, CA, USA
e-mail: gowen@nps.edu

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_36,

is devoted to giving two axiomatic characterizations of the Shapley rule on the space of all PBPs, and also on several subsets of interest: among them, the domain of the proportional rule, the open positive and negative orthants, and the cone of (strict) superadditive PBPs. In Sect. 6, we present a criticism on the proportional rule, referring mainly to its inconsistency in related cost-saving problems and added costs problems. Section 7 collects the conclusions of the work.

2 Pure Bargaining Problems and Sharing Rules

Let $N = \{1, 2, \ldots, n\}$ (with $n \geq 2$) be a set of agents and assume that there are given: (a) a set of utilities $u_1, u_2, \ldots, u_n$ available to the agents individually and (b) a total utility u_N that, alternatively, the agents can jointly get if all of them agree—utilities denoting costs will be represented by negative numbers. A vector $u = (u_1, u_2, \ldots, u_n | u_N)$ collects all this information and we will say that it represents a *pure bargaining problem* (*PBP*, in the sequel) on N. The *surplus* of u is defined as

$$\Delta(u) = u_N - \sum_{j \in N} u_j.$$

The problem consists in dividing u_N among the agents in a rational way, i.e. in such a manner that all of them should agree and feel (more or less) satisfied with the outcome. Of course, the individual utilities $u_1, u_2, \ldots, u_n$ should be taken into account. The transferable utility assumption will mean that any vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ with $x_1 + x_2 + \cdots + x_n = u_N$ is feasible if the n agents agree.

Example 2.1 (*A cost allocation problem*) Assume that three towns, A, B and C, wish to get some kind of supply (electricity, water, gas) from a supplier S. The locations are $A(2,2)$, $B(-2,2)$, $C(-2,-2)$ and $S(2,0)$, the distances given in kilometers (see Fig. 1). The connection cost amounts to 100 monetary units per km.

For individual connections, the supplier offers lines SA, SB and SC. For A, B and C together, the offer consists in using SA, SO, OB and OC. The question is how to share the joint connection cost. Then we have a (rounded) *cost PBP* $u^c = (-200, -448, -448 | -966)$ that describes the individual and joint costs and is defined on $N = \{1, 2, 3\}$, where 1 is A, 2 is B and 3 is C. Assume that the three towns sign a joint contract with the supplier. How should they share the total cost of 966?

An equivalent approach is obtained when considering the *saving PBP* $u^s = (0, 0, 0 | 130)$, which gives the savings derived from agreeing or not the joint contract. Now the question is: how should the three towns share the net savings of 130 for a whole contract? Of course, there should exist a *consistent* solution for both cost and saving (related) problems.

Let $E_{n+1} = \mathbb{R}^n \times \mathbb{R}$ denote the $(n+1)$-dimensional vector space formed by all PBPs on N. In order to deal with, and solve, all possible PBPs on N, one should

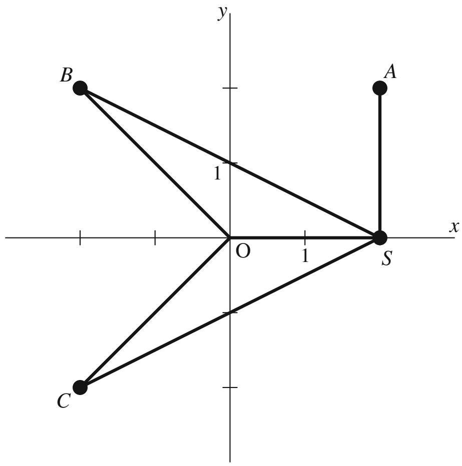


Fig. 1 Towns and supplier positions

look for a *sharing rule*, i.e. a function $f : E_{n+1} \longrightarrow \mathbb{R}^n$. Given $u \in E_{n+1}$, for each $i \in N$ the i-coordinate $f_i[u]$ will provide the share of u_N that corresponds to agent i according to f. Of course, there are infinitely many such functions: for example, $f_1[u] = u_N$ and $f_i[u] = 0$ for $i \neq 1$ would define one of them. More interesting ideas are given by the proportional rule, often used in practice and denoted here by π, and the equal surplus sharing rule, denoted by ε.

Definition 2.2 (a) The *proportional rule* π is defined by

$$\pi_i[u] = \frac{u_i}{u_1 + u_2 + \cdots + u_n} u_N \qquad \text{for each } i \in N. \tag{1}$$

For further purposes, we notice that this expression can be transformed as follows:

$$\pi_i[u] = \frac{u_i}{u_1 + u_2 + \cdots + u_n} u_N = u_i + \frac{u_i\, u_N}{\sum_{j \in N} u_j} - u_i = u_i + \frac{u_i}{\sum_{j \in N} u_j} \Delta(u). \tag{2}$$

However, a main problem is that the domain of the proportional rule is not E_{n+1} but the subset

$$E^{\pi}_{n+1} = \{u \in E_{n+1} : u_1 + u_2 + \cdots + u_n \neq 0\}, \tag{3}$$

that is, the complement of a hyperplane.

(b) Instead of this, the *equal surplus sharing rule* ε is defined by

$$\varepsilon_i[u] = u_i + \frac{\Delta(u)}{n} \qquad \text{for each } i \in N,$$

so that its domain is the entire space E_{n+1} without restriction.

Remark 2.3 Kalai (1977) discusses what he calls a "proportional rule". This is by no means the same as the proportional rule that we discuss here. In fact, both Kalai's rule and our π are proportional in the sense that the excess utility $\Delta(u)$ is divided among the agents in proportion to something. In Kalai's rule, the division is in proportion to an exogenously given point $\mathbf{p} = (p_1, p_2, \ldots, p_n)$. Thus each agent receives the amount $u_i + \lambda p_i$, where λ is as large as possible, subject to feasibility.

There is yet another "proportional rule", this one due to Kalai and Smorodinsky (1975). In this latter rule, the division is in proportion to the maximum gain ξ_i that each agent could eventually hope for, assuming that all the other agents receive their conflict utility u_j, i.e.

$$\xi_i = \max\{x_i : (u_1, u_2, \ldots, u_{i-1}, x_i, u_{i+1}, \ldots, u_n) \text{ is feasible}\} - u_i.$$

Each player then receives the quantity $u_i + \lambda\xi_i$, where λ is as large as possible, subject of course to feasibility.

Note that, in our case (with transferable utility), this $\xi_i = \Delta(u)$ for each i. Thus, in our domain E_{n+1}, the Kalai–Smorodinsky rule coincides, not with our proportional rule π, but rather with the equal-sharing rule ε.

Example 2.4 (*A basis of PBPs*) Special elementary PBPs will be useful in Sect. 5. They are defined as follows:

- $u_0^k = (1, \ldots, 1, \overset{k}{\smile}{2}, 1, \ldots, 1|n+1)$ for $k = 1, 2, \ldots, n$ and
- $u_0^N = (1, 1, \ldots, 1|n+1)$.

It is not difficult to check that $\mathcal{B}_0 = \{u_0^1, u_0^2, \ldots, u_0^n, u_0^N\}$ is a basis for E_{n+1}. Hence each $u \in E_{n+1}$ can be uniquely written as a linear combination

$$u = x_1 u_0^1 + x_2 u_0^2 + \cdots + x_n u_0^n + x_N u_0^N.$$

If $u = (u_1, u_2, \ldots, u_n|u_N)$, its components relative to $\mathcal{B}_0$ are easily found. By introducing the surplus we obtain

$$u = \sum_{k=1}^{n} \left(u_k - \frac{u_N}{n+1}\right) u_0^k + \Delta(u) u_0^N. \tag{4}$$

3 Closures and Quasi-Additive Games

The Shapley value (Shapley 1953; see also Roth (1988) or Owen (1995)), denoted here by φ, cannot be directly applied to PBPs as a sharing rule. We will therefore associate a TU (i.e., side-payment) cooperative game with each PBP in a natural way. Let $\mathcal{G}_N$ be the vector space of all cooperative TU games with N as set of players and let us define a map $\sigma : E_{n+1} \longrightarrow \mathcal{G}_N$ as follows. If $u = (u_1, u_2, \ldots, u_n | u_N)$ then $\overline{u} = \sigma(u)$ is given by

$$\overline{u}(S) = \begin{cases} \sum_{i \in S} u_i & \text{if} \quad S \neq N, \\ u_N & \text{if} \quad S = N. \end{cases}$$

The idea behind this definition is simple. Since, given a PBP u, nothing is known about the utility available to each intermediate coalition $S \subset N$ with $|S| > 1$, a reasonable assumption is that such a coalition can get the sum of the individual utilities of its members (cf., however, Remark 4.2). Game $\overline{u}$ will be called the *closure* of u. It is not difficult to verify the following properties of σ:

- σ is a linear map.
- σ is one-to-one, i.e. $\ker\{\sigma\} = \{0\}$.

Let us recall that a cooperative game v is *additive* iff $v(S) = \sum_{i \in S} v(\{i\})$ for all $S \subseteq N$. If we drop this condition just for $S = N$ and give the name *quasi-additive* to the games that fulfill it for all $S \subset N$, it follows that these games precisely form the image set $\mathrm{Im}(\sigma)$, and hence a game is quasi-additive iff it is the closure of a PBP, which is unique. The dimension of the subspace of quasi-additive games is $n + 1$. (If $n = 2$ then σ is onto and therefore *any* cooperative 2-person game is the closure of a PBP.)

As σ is an embedding of E_{n+1} into $\mathcal{G}_N$ (cf. also Remark 4.3), reasonable restrictions for games can be adapted to PBPs after identifying each PBP with the corresponding closure. Then, we call to a PBP $u \in E_{n+1}$

- *additive* (or *inessential*) iff $u_1 + u_2 + \cdots + u_n = u_N$, i.e. iff $\Delta(u) = 0$;
- *superadditive* (strictly) iff $u_1 + u_2 + \cdots + u_n < u_N$, i.e. iff $\Delta(u) > 0$;
- *symmetric* iff $u_i = u_j$ for all $i, j \in N$;
- *positive* iff $u_1, u_2, \ldots, u_n, u_N > 0$;
- *negative* iff $u_1, u_2, \ldots, u_n, u_N < 0$;
- *nonpositive* iff $u_1, u_2, \ldots, u_n, u_N \leq 0$; and
- *nonnegative* iff $u_1, u_2, \ldots, u_n, u_N \geq 0$.

Thus, in Example 2.4 $u_0^1, u_0^2, \ldots, u_0^n$ are additive but not symmetric whereas u_0^N is superadditive and symmetric, and all of them are positive and belong to E_{n+1}^{π}, the domain of π.

4 Core and the Shapley Rule

Now, let us first apply the Shapley value to any quasi-additive game, i.e. to $\overline{u}$ for any $u \in E_{n+1}$, and get an explicit formula. For each $i \in N$ we have

$$\varphi_i[\overline{u}] = \sum_{S \ni i} \frac{(s-1)!(n-s)!}{n!} [\overline{u}(S) - \overline{u}(S \backslash \{i\})] = \sum_{\substack{S \ni i \\ S \neq N}} \frac{(s-1)!(n-s)!}{n!} u_i + \frac{1}{n} \left[u_N - \sum_{j \neq i} u_j \right]$$

$$= u_i \sum_{S \ni i} \frac{(s-1)!(n-s)!}{n!} + \frac{1}{n} \left[u_N - \sum_{j \in N} u_j \right] = u_i + \frac{\Delta(u)}{n},$$

since, for all $i \in N$,

$$\sum_{S \ni i} \frac{(s-1)!(n-s)!}{n!} = 1.$$

A quasi-additive game $\overline{u} = \sigma(u)$ is convex (Shapley 1971), that is, satisfies

$$\overline{u}(S) + \overline{u}(T) \leq \overline{u}(S \cap T) + \overline{u}(S \cup T) \quad \text{for all } S, T \subseteq N,$$

iff $\Delta(u) \geq 0$. The core (Gillies 1953), given for a general cooperative game v by

$$C(v) = \{x = (x_1, x_2, \ldots, x_n) : \sum_{i \in N} x_i = v(N) \text{ and } \sum_{i \in S} x_i \geq v(S) \text{ for all } S \subset N\},$$

takes here the much simpler form

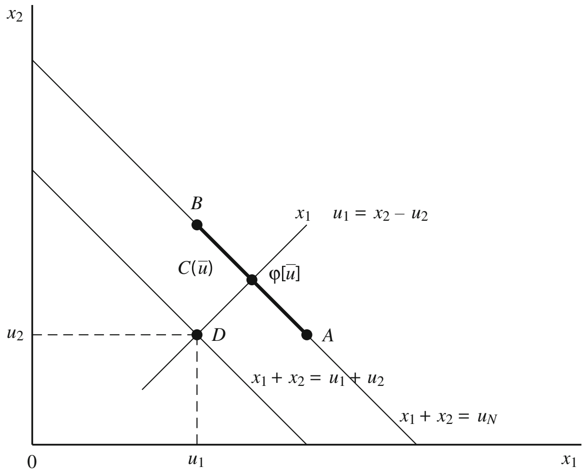
$$C(\overline{u}) = \{x = (x_1, x_2, \ldots, x_n) : \sum_{i \in N} x_i = u_N \quad \text{and } x_i \geq u_i \quad \text{for all } i = 1, 2, , n\}.$$

$C(\overline{u})$ is nonempty iff $\Delta(u) \geq 0$, and then the Shapley value $\varphi[\overline{u}] \in C(\overline{u})$ according to Shapley (1971). In fact, for a quasi-additive game the nonnegativity of the surplus is equivalent to, and not only sufficient for, the nonemptiness of the core.

Figure 2 describes the core geometrically for $n = 2$. In the interesting case, when $\Delta(u) > 0$, the core is the closed segment AB on the line $x_1 + x_2 = u_N$. The Shapley value $\varphi[\overline{u}]$ is the intersection of this line with $x_1 - u_1 = x_2 - u_2$, the orthogonal line from the *disagreement point D*. In other words, the Shapley value is the orthogonal projection of the disagreement point onto the core. As a limiting case, if $\Delta(u) = 0$ then the line $x_1 + x_2 = u_N$ reduces to the line $x_1 + x_2 = u_1 + u_2$, A and B coincide with D, and hence the core reduces to this disagreement point, which coincides with the Shapley value of the game. Finally, if $\Delta(u) < 0$ the core of $\overline{u}$ is empty.

The generalization of these ideas to arbitrary n is straightforward. The disagreement point D is given by $x_1 = u_1,\ x_2 = u_2,\ \ldots,\ x_n = u_n$. The core is the simplex defined by $x_1 \geq u_1,\ x_2 \geq u_2,\ \ldots,\ x_n \geq u_n$ in the hyperplane $x_1 + x_2 + \cdots + x_n = u_N$. It becomes empty if $\Delta(u) < 0$ and reduces to the disagreement point if $\Delta(u) = 0$. Otherwise, that is, whenever $\Delta(u) > 0$, *the Shapley value* $\varphi[\overline{u}]$ *is the orthogonal projection of the disagreement point onto the core*, i.e. the intersection of the core with the orthogonal line $x_1 - u_1 = x_2 - u_2 = \ldots = x_n - u_n$.

Fig. 2 Core and Shapley value of a superadditive quasi-additive game $\overline{u}$ for $n = 2$



We find here, thus, a particular case of Nash's classical bargaining problem (Nash 1950). The feasible set S is defined by $\sum_{i \in N} x_i \leq u_N$, the Pareto frontier is given by $\sum_{i \in N} x_i = u_N$, and the disagreement point is $D = (u_1, u_2, \ldots, u_n)$, which may lie above the Pareto frontier (just in case that $\Delta(u) < 0$). Moreover, $\varphi[\overline{u}]$ coincides with the Nash solution. It should be noticed that, for these games, $\varphi[\overline{u}]$ also coincides with the (Kalai and Smorodinsky 1975) solution. (Of course, both Nash and Kalai–Smorodinsky solutions are defined for 2-person problems only, but they generalize directly to n-person PBPs.)

We are now ready to introduce the Shapley rule for PBPs.

Definition 4.1 By setting

$$\overline{\varphi}[u] = \varphi[\overline{u}] \quad \text{for all } u \in E_{n+1} \tag{5}$$

we obtain a function $\overline{\varphi} : E_{n+1} \longrightarrow \mathbb{R}^n$. Function $\overline{\varphi}$ will be called the *Shapley rule* (for PBPs). It is given by

$$\overline{\varphi}_i[u] = u_i + \frac{\Delta(u)}{n} \quad \text{for each } i \in N \text{ and each } u \in E_{n+1}, \tag{6}$$

and hence it solves each PBP in the following way: (a) first, each agent is allocated his individual utility; (b) once this has been done, the remaining utility is equally shared among all agents.

Notice that $\overline{\varphi}$ is linear, coincides with ε, the equal surplus sharing rule introduced in Definition 2.2, and is reminiscent of the CIS-value introduced in Driessen and Funaki (1991). Besides, our rule can be compared to the (Aadland and Kolpin 1998) egalitarian solution. In fact, the two coincide if $u_i = 0$ for all $i \in N$. Otherwise, our rule has the property that, for *congruent* PBPs (in the sense of

Neumann, von and Morgenstern (1944) for games), it is covariant with respect to utility transformations: if there are $\alpha, \beta_1, \beta_2, \ldots, \beta_n \in \mathbb{R}$ such that

$$u_i = \alpha v_i + \beta_i \quad \text{for each } i \in N \quad \text{and} \quad u_N = \alpha v_N + \sum_{i \in N} \beta_i,$$

then $\overline{\varphi}_i[u] = \alpha \overline{\varphi}_i[v] + \beta_i$.

The Shapley rule shows an "egalitarian flavor" in the sense of Brink, van den and Funaki (2009). Indeed, from Eq. (6) it follows that the Shapley rule is a mixture consisting of a "competitive" component, which rewards each agent according to the individual utility, and a "solidarity" component that treats all agents equally when sharing the surplus. Thus it satisfies standardness for two–agent PBPs in the sense of Hart and Mas-Colell (1988) and Hart and Mas-Colell (1989).

Remark 4.2 Our closure definition might seem rather pessimistic. For each $S \subset N$, the aspiration of the agents of S as a team in $\overline{u}$ is reduced to obtaining the sum of their individual utilities. Thus, if $\Delta(u) > 0$ they implicitly leave $u_N - \sum_{i \in S} u_i > \sum_{j \notin S} u_j$ to the remaining agents. Notice, however, that if $\Delta(u) < 0$ the definition becomes optimistic since what they leave in this case is less than $\sum_{j \notin S} u_j$.

An alternative closure notion could be adopted by following the approach suggested in e.g. Bergantiños and Vidal–Puga (2007). Let us define a second closure $\overline{u}^*$ of any PBP u by

$$\overline{u}^*(S) = \begin{cases} 0 & \text{if} \quad S = \emptyset, \\ u_N - \sum_{j \notin S} u_j & \text{if} \quad S \neq \emptyset. \end{cases}$$

This second closure notion is optimistic if $\Delta(u) > 0$ because it implicitly assumes that the agents j not in S will be satisfied by receiving just u_j each, and the remaining utility will therefore be allocated to the members of S. But, instead, if $\Delta(u) < 0$ then this closure notion is pessimistic.

We have $\overline{u}^*(N) = u_N$ but, in general, $\overline{u}^*(\{i\}) = u_i$ does not hold for all $i \in N$. The map $\tau : E_{n+1} \longrightarrow \mathcal{G}_N$ defined by $\tau(u) = \overline{u}^*$ is linear and one-to-one but, of course, $\mathrm{Im}(\tau) \neq \mathrm{Im}(\sigma)$: if v_S stands for the unanimity game attached to a non-empty coalition $S \subseteq N$, $\mathrm{Im}(\sigma)$ is the subspace spanned by $v_{\{1\}}, v_{\{2\}}, \ldots, v_{\{n\}}$ and v_N, whereas $\mathrm{Im}(\tau)$ is the subspace spanned by $v_{\{1\}}, v_{\{2\}}, \ldots, v_{\{n\}}$ and $v = \sum_{S \neq \emptyset, N} (-1)^s v_S$, where $s = |S|$.

Let us see which would be the Shapley rule associated with this second closure notion. First, for any PBP u, $\delta = \overline{u}^* - \overline{u}$ is clearly given by

$$\delta(S) = \begin{cases} \Delta(u) & \text{if} \quad \emptyset \subset S \subset N, \\ 0 & \text{if } S = \emptyset \ \text{ or } S = N. \end{cases}$$

As δ is a symmetric game and $\delta(N)=0$, it follows that $\varphi[\overline{u}^*-\overline{u}]=\varphi[\delta]=0$, and hence $\varphi[\overline{u}^*]=\varphi[\overline{u}]$ for all $u\in E_{n+1}$. The conclusion is that the two Shapley rules (Definition 4.1 and this remark, resp.) coincide.

Remark 4.3 The passage from φ on games to $\overline{\varphi}$ on PBPs is in some manner antiparallel to the passing from π on PBPs to a value $\overline{\pi}$ on games. Indeed, let $\mathcal{G}_N^\pi$ denote the set of games v that fulfill the condition

$$v(\{1\})+v(\{2\})+\cdots+v(\{n\})\neq 0.$$

Let us define $\psi:\mathcal{G}_N^\pi\longrightarrow E_{n+1}$ by $\psi(v)=\overline{v}$, where $\overline{v}_i=v(\{i\})$ for each $i\in N$ and $\overline{v}_N=v(N)$. This gives us a linear map ψ with $\operatorname{Im}(\psi)=E_{n+1}^\pi$ and such that $\psi\circ\sigma=id$ on E_{n+1} but, in general, $\sigma\circ\psi\neq id$ on $\mathcal{G}_N^\pi$. By setting

$$\overline{\pi}[v]=\pi[\overline{v}]\quad\text{ for all } v\in\mathcal{G}_N^\pi$$

we have a function $\overline{\pi}:\mathcal{G}_N^\pi\longrightarrow\mathbb{R}^n$ given by

$$\overline{\pi}_i[v]=\frac{v(\{i\})}{v(\{1\})+v(\{2\})+\cdots+v(\{n\})}v(N)\quad\text{ for each } i\in N \text{ and each } v\in\mathcal{G}_N^\pi.$$

The so-called *proportional value* $\overline{\pi}$ has been studied in Ortmann (2000) as a value on the domains of *positive* (resp., *negative*) games, that is, games v where $v(S)>0$ (resp., $v(S)<0$) for all nonempty $S\subseteq N$.

To close this section we determine the set of PBPs where the Shapley rule and the proportional rule coincide and discuss individual rationality.

Proposition 4.4 *The Shapley rule and the proportional rule coincide on a PBP* $u\in E_{n+1}^\pi$ iff u *is additive or symmetric.*

Proof ($\Leftarrow$) If $u\in E_{n+1}^\pi$ is additive (resp., symmetric), from Eqs. (1) and (6) it readily follows that, for all $i\in N$,

$$\pi_i[u]=u_i=\overline{\varphi}_i[u]\quad\left(\text{ resp.,}\quad \pi_i[u]=\frac{u_N}{n}=\overline{\varphi}_i[u]\right).$$

($\Rightarrow$) Assume that $\pi[u]=\overline{\varphi}[u]$ for some $u\in E_{n+1}^\pi$. Then, using Eq. (2),

$$u_i+\frac{u_i}{\sum_{j\in N}u_j}\Delta(u)=u_i+\frac{\Delta(u)}{n}\quad\text{ for all } i\in N.$$

Canceling the common initial term yields

$$\left(\frac{u_i}{\sum_{j\in N}u_j}-\frac{1}{n}\right)\Delta(u)=0\quad\text{ for all } i\in N.$$

Then, either:

(a) $\Delta(u) = 0$ and the PBP u is additive, or
(b) $u_i = \frac{1}{n}\sum_{j \in N} u_j$ for all $i \in N$ and the PBP u is symmetric. □

Remark 4.5 For any sharing rule f, the property of *individual rationality* states that

$$f_i[u] \geq u_i \quad \text{for all } i \in N \text{ and all } u \in E_{n+1}.$$

It is easy to verify that the Shapley rule satisfies this property just in the domain of all additive or superadditive PBPs. Indeed, using Eq. (6) it follows that, for all $i \in N$,

$$\overline{\varphi}_i[u] \geq u_i \quad \text{iff} \quad \Delta(u) \geq 0.$$

In $u = (-2, 1|1)$, a troubling example suggested by a reviewer where $\Delta(u) = 2$ and hence superadditivity holds, we get $\overline{\varphi}[u] = (-1, 2)$, an individually rational and quite reasonable result. Instead, the proportional rule gives $\pi[u] = (2, -1)$, a completely counterintuitive output. The domain where the proportional rule satisfies individual rationality, not difficult but cumbersome to describe, covers only a fraction of the set of superadditive PBPs.

5 Axiomatic Characterizations of the Shapley Rule

When looking for a function $f : E_{n+1} \longrightarrow \mathbb{R}^n$, some reasonable properties should be imposed. To state a set of them, we previously define dummy agent, null agent, and symmetric agents in a PBP. Agent $i \in N$ is a *dummy* in a PBP u iff $u_N = u_i + \sum_{j \neq i} u_j$, and *null* if, moreover, $u_i = 0$. Agents $i, j \in N$ are *symmetric* in a PBP u iff $u_i = u_j$.

For instance, in Example 2.4 all agents are dummies but not null in each u_0^k, all $i \neq k$ are symmetric in u_0^k, and all i are symmetric in u_0^N. A PBP with a dummy agent must be additive and therefore all agents are dummies. Conversely, if a PBP is additive (i.e., inessential) then all agents are dummies.

5.1 Main Theorem

Let us consider the following properties, stated (for a function f defined) on E_{n+1}:

(i) *Efficiency*: $\sum_{i \in N} f_i[u] = u_N$ for every u.
(ii) *Dummy agent property*: if i is a dummy in u then $f_i[u] = u_i$.
(iii) *Symmetry*: if i and j are symmetric in u then $f_i[u] = f_j[u]$.
(iv) *Additivity*: $f[u + v] = f[u] + f[v]$ for all u, v.

These properties deserve to be called "axioms" because of their elegant simplicity. It is hard to claim that they are not compelling. Efficiency means that the agents are going to share the total amount available to them. The dummy agent property essentially says that if a PBP is additive then each agent should receive his individual utility. Symmetry establishes that two agents that are equally powerful individually should receive the same payoff. Finally, additivity implies that the allocation in a sum of PBPs must coincide with the sum of allocations in each PBP.

The question is the following: is there some function satisfying properties (i)–(iv)? If so, is it unique? The positive answers are given in the following result.

Theorem 5.1 (*First main axiomatic characterization of the Shapley rule*) *There is one and only one function $f : E_{n+1} \longrightarrow \mathbb{R}^n$ that satisfies properties (i)–(iv). It is the Shapley rule $\overline{\varphi}$.*

Proof (Existence) It suffices to show that the Shapley rule satisfies (i)–(iv), and this follows at once from Eq. (6).

(Uniqueness) We shall see that if f satisfies (i)–(iv) then it is completely determined. To this end, we will use the basis $\mathcal{B}_0$ of E_{n+1} introduced in Example 2.4. For any $\lambda \in \mathbb{R}$, from (i) and (ii) it follows that, for any $i, k \in N$,

$$f_i[\lambda u_0^k] = \begin{cases} 2\lambda & \text{if } i = k, \\ \lambda & \text{if } i \neq k, \end{cases}$$

and, from (i) and (iii),

$$f_i[\lambda u_0^N] = \lambda \frac{n+1}{n} \qquad \text{for all } i \in N.$$

Then, for any $i \in N$ and any $u \in E_{n+1}$, from (iv) and using Eq. (4) we obtain

$$f_i[u] = 2\left(u_i - \frac{u_N}{n+1}\right) + \sum_{k \neq i} \left(u_k - \frac{u_N}{n+1}\right) + \Delta(u)\frac{n+1}{n} = u_i + \frac{\Delta(u)}{n} = \overline{\varphi}_i[u],$$

so that $f = \overline{\varphi}$.

□

Remark 5.2 (a) As was mentioned after introducing it at the beginning of this section, the dummy agent property (ii) is clearly equivalent to the following property:

(ii)' *Inessential problem*: if u is additive then $f_i[u] = u_i$ for all $i \in N$.

Moreover, the dummy agent property (ii) is a generalization of the following one:

(ii)" *Null agent property*: if i is null in u then $f_i[u] = 0$.

As the Shapley rule satisfies it, (ii)" could replace (ii) in Theorem 5.1. The uniqueness part of the proof should then be modified accordingly by replacing the basis $\mathcal{B}_0$, none of whose members possesses null agents, with e.g. the alternative basis $\mathcal{B}_a = \{u_a^1, u_a^2, \ldots, u_a^n, u_a^N\}$ given by

- $u_a^k = (0, \ldots, 0, \overset{k}{\overbrace{1}}, 0, \ldots, 0|1)$ for $k = 1, 2, \ldots, n$ and
- $u_a^N = (1, 1, \ldots, 1|n + 1)$.

(b) Readers aware of Shapley's seminal work for cooperative games (Shapley 1953) will not be greatly surprised by Theorem 5.1. It may be noticed that it is equivalent to an axiomatic characterization of (the restriction of) the Shapley value on the subspace of quasi-additive games, since the notions of dummy agent, null agent, and symmetric agents in a PBP correspond, *mutatis mutandis*, to the respective notions of dummy player, null player, and symmetric players in the closure of the given PBP. In fact, in the existence proof of Theorem 5.1, all properties would follow from Eq. (5), the additivity of σ, and the corresponding properties of the classical Shapley value on games. A similar remark would apply to Theorem 5.4 with regard to Young's work (Young 1985).

5.2 Other Domains

Several subsets of E_{n+1} deserve special attention, and it would be therefore of interest to have axiomatic characterizations for (the restriction of) the Shapley rule on each one of these domains. So as not to enlarge the analysis too much, we will restrict it to the following ones:

- The domain of the proportional rule: $E_{n+1}^{\pi} = \{u \in E_{n+1} : u_1 + u_2 + \cdots + u_n \neq 0\}$. We wish to contrast below (Sect. 6.3) the Shapley rule and the proportional rule strictly in this domain, in order to give "all advantages" to π (if any) in our discussion.
- The open orthant of positive PBPs: $E_{n+1}^{++} = \{u \in E_{n+1} : u_1, u_2, \ldots, u_n, u_N > 0\}$. σ maps this into the subset of positive games considered in Ortmann (2000). A similar reason applies to the open orthant of negative PBPs: $E_{n+1}^{--} = \{u \in E_{n+1} : u_1, u_2, , u_n, u_N < 0\}$.
- The open cone of superadditive PBPs: $E_{n+1}^{sa} = \{u \in E_{n+1} : u_1 + u_2 + \cdots + u_n < u_N\}$. These PBPs, where the surplus is $\Delta(u) > 0$, are the most interesting ones since in each one of them there is something to gain by cooperation.
- The intersection of the cone of superadditive PBPs with the closed orthant of nonnegative PBPs: $E_{n+1}^{+} = \{u \in E_{n+1} : u_1, u_2, \ldots, u_n, u_N \geq 0\}$. And also the intersection of the cone with the closed orthant of nonpositive PBPs: $E_{n+1}^{-} = \{u \in E_{n+1} : u_1, u_2, \ldots, u_n, u_N \leq 0\}$. The former includes all profit PBPs, while the latter includes all cost PBPs. By combining with the positivity of the

surplus in both cases we obtain the two most appealing types of PBPs in practice.

If E denotes any of the subsets of E_{n+1} mentioned just above, properties (i)–(iv) make sense for $f : E \longrightarrow \mathbb{R}^n$ if we state them only for $u, v \in E$. The sole exception is property (iv) for E^{π}_{n+1} since it is the only one of these domains not closed under addition of PBPs. Therefore in this case we will assume that the property is:

(iv) *Additivity*: if $u, v, u + v \in E^{\pi}_{n+1}$ then $f[u + v] = f[u] + f[v]$.

Theorem 5.3 (*Additional axiomatic characterizations of the Shapley rule*) *For either*

$$E = E^{\pi}_{n+1}, \quad E = E^{++}_{n+1}, \quad E = E^{--}_{n+1}, \quad E = E^{sa}_{n+1}, \quad E = E^{+}_{n+1} \cap E^{sa}_{n+1} \quad \text{or} \quad E = E^{-}_{n+1} \cap E^{sa}_{n+1},$$

there is one and only one function $f : E \longrightarrow \mathbb{R}^n$ that satisfies properties (i)–(iv). In all cases it is (the restriction of) the Shapley rule $\overline{\varphi}$.

Proof (Existence) In all cases, it derives from the existence part of Theorem 5.1.

(Uniqueness) In all cases again the proofs are similar to that of uniqueness in Theorem 5.1, but some points must be handled with care. We shall see that if f satisfies (i)–(iv) on E then it is completely determined. As a matter of notation, we will use $\overline{N} = \{1, 2, \ldots, n, N\}$ as a set of indices.

(a) $E = E^{\pi}_{n+1}$. As in Theorem 5.1, from properties (i)–(iii) it follows that f is determined on any nonzero multiple of each member of the basis $\mathcal{B}_0$, all of which belong to E^{π}_{n+1}. Now, let $u \in E^{\pi}_{n+1}$, and hence $u \neq 0$. By removing all terms of Eq. (4) whose coefficient vanishes, we get

$$u = \sum_{j \in S} \alpha_j u^j_0 - \sum_{k \in T} (-\beta_k) u^k_0,$$

where $S = \{j \in \overline{N} : \alpha_j > 0\}$, $T = \{k \in \overline{N} : \beta_k < 0\}$ and $S \cup T \neq \emptyset$. Then,

$$u + \sum_{k \in T} (-\beta_k) u^k_0 = \sum_{j \in S} \alpha_j u^j_0.$$

Since, for all $\gamma > 0$ and all $i \in \overline{N}$, all PBPs of the form γu^i_0, as well as any sum of them, are in E^{π}_{n+1}, a repeated application of (iv) yields

$$f[u] + \sum_{k \in T} f[-\beta_k u^k_0] = \sum_{j \in S} f[\alpha_j u^j_0],$$

and hence f is also determined on any $u \in E^{\pi}_{n+1}$.

The proofs in the remaining cases are completely analogous if the bases indicated below for each one of them are used.

(b) $E = E^{++}_{n+1}$. Suitable basis: $\mathcal{B}_0$ again, since all its members are positive.

(c) $E = E^{--}_{n+1}$. Suitable basis: $\mathcal{B}_b = \{u_b^1, u_b^2, \ldots, u_b^n, u_b^N\}$, where $u_b^k = -u_0^k$ for $k = 1, 2, \ldots, n, N$.

(d) $E = E^{sa}_{n+1}$. Suitable basis: $\mathcal{B}_c = \{u_c^1, u_c^2, \ldots, u_c^n, u_c^N\}$, where

$$u_c^k = (1, \ldots, 1, \overset{k}{\smile}{2}, 1, , 1|n+2) \quad \text{for} \quad k = 1, 2, \ldots, n \quad \text{and} \quad u_c^N = (1, 1, \ldots, 1|n+1).$$

(e) $E = E^{+}_{n+1} \cap E^{sa}_{n+1}$. Suitable basis: $\mathcal{B}_c$ again.

(f) $E = E^{-}_{n+1} \cap E^{sa}_{n+1}$. Suitable basis: $\mathcal{B}_d = \{u_d^1, u_d^2, , u_d^n, u_d^N\}$, where

$$u_d^k = (-1, \ldots, -1, \overset{k}{\smile}{-2}, -1, \ldots, -1| -n) \quad \text{for} \quad k = 1, 2, \ldots, n \quad \text{and}$$
$$u_d^N = (-1, -1, \ldots, -1| -n+1). \qquad \square$$

5.3 *Discussing Monotonicity*

In the literature on cooperative games, several monotonicity conditions have been suggested for solution concepts. Here we will recall some of the most relevant ones and will adapt them to the PBP setup, i.e., for sharing rules.

Let $\bar{u}, \bar{v} \in \mathcal{G}_N$ and $g : \mathcal{G}_N \longrightarrow \mathbb{R}^n$ be a solution concept. *Coalitional monotonicity* (Shubik 1962) states that if $\bar{u}(T) \geq \bar{v}(T)$ for some $T \subseteq N$ and $\bar{u}(S) = \bar{v}(S)$ for all $S \neq T$ then $g_i[\bar{u}] \geq g_i[\bar{v}]$ for all $i \in T$. In the particular case where $T = N$ we obtain *aggregate monotonicity* (Megiddo 1974). *Strong monotonicity* (Young 1985) refers to marginal contributions and states that if $\bar{u}(S) - \bar{u}(S \backslash \{i\}) \geq \bar{v}(S) - \bar{v}(S \backslash \{i\})$ for all $S \subseteq N$ then $g_i[\bar{u}] \geq g_i \bar{v}]$ for that i. The Shapley value satisfies all these conditions (see Young 1985). However, on quasi-additive games, coalitional monotonicity makes sense only for $T = N$, thus reducing to aggregate monotonicity which, in turn, becomes a consequence of strong monotonicity. Then we will translate to PBPs only this last property.

From a different approach, new interesting monotonicity conditions, among which we find again strong monotonicity, have been proposed in Carreras and Freixas (2000). They are based on: (a) the *desirability relation* D, introduced in Isbell (1956) for the players of any game $\bar{u}$ and given by

$$iDj \text{ in } \bar{u} \qquad \text{iff} \quad \bar{u}(S \cup \{i\}) - \bar{u}(S) \geq \bar{u}(S \cup \{j\}) - \bar{u}(S) \text{ for all } S \subseteq N \backslash \{i, j\},$$

which compares *the positions of two players in a common game*; and (b) a similar relation B, introduced in Carreras and Freixas (2000) and given by

$$\bar{u} \, B \, \bar{v} \text{ for } i \qquad \text{iff} \qquad \bar{u}(S \cup \{i\}) - \bar{u}(S) \geq \bar{v}(S \cup \{i\}) - \bar{v}(S) \quad \text{for all } S \subseteq N \backslash \{i\},$$

which compares *the positions of a common player in two games*. These conditions are:

- *monotonicity*: if iDj in $\overline{u}$ then $g_i[\overline{u}] \geq g_j[\overline{u}]$
- *strict monotonicity*: if iDj and $j\not{D}i$ in $\overline{u}$ then $g_i[\overline{u}] > g_j[\overline{u}]$
- *strong monotonicity*: if $\overline{u}\ B\ \overline{v}$ for i then $g_i[\overline{u}] \geq g_i[\overline{v}]$
- *strict strong monotonicity*: if $\overline{u}\ B\ \overline{v}$ and $\overline{v}\ \not{B}\,\overline{u}$ for i then $g_i[\overline{u}] > g_i[\overline{v}]$.

The Shapley value satisfies all these conditions (see Carreras and Freixas 2000). Moreover, for any quasi-additive game $\overline{u} = \sigma(u)$, where $u = (u_1, u_2, \ldots, u_n | u_N) \in E_{n+1}$, the marginal contribution to $S \subseteq N$ of a player $i \in S$ in $\overline{u}$ is given by

$$\overline{u}(S) - \overline{u}(S\backslash\{i\}) = \begin{cases} u_i & \text{if} \quad S \neq N, \\ u_N - \sum\limits_{j \neq i} u_j & \text{if} \quad S = N, \end{cases}\}$$

and it follows that

- iDj in $\overline{u}$ iff $u_i \geq u_j$
- $\overline{u}\ B\,\overline{v}$ for i iff $u_i \geq v_i$ and $\Delta(u) + u_i \geq \Delta(v) + v_i$.

Thus, relations D and B, as well as the four monotonicity conditions stated above, make sense in PBPs (just replacing each quasi-additive game $\overline{u}$ with the corresponding PBP u given by σ^{-1}), and the Shapley rule $\overline{\varphi}$ satisfies all these conditions.

Once within the PBP framework, we obtain a new main axiomatic characterization of the Shapley rule on E_{n+1} that is quite different from Theorem 5.1 and is reminiscent of Young (1985) characterization of the Shapley value without using additivity.

Theorem 5.4 (*Second main axiomatic characterization of the Shapley rule*) *There is one and only one function* $f : E_{n+1} \longrightarrow \mathbb{R}^n$ *that satisfies efficiency, symmetry and strong monotonicity. It is the Shapley rule* $\overline{\varphi}$.

Proof (Existence) It follows from the existence proof in Theorem 5.1 and the fact already stated that the Shapley rule satisfies strong monotonicity.

(Uniqueness) We shall see that if f satisfies these three properties then it is completely determined. The proof will consist of several steps.

1. Let $0 = (0, 0, \ldots, 0|0)$ be the null PBP. By efficiency and symmetry it follows that $f_i[0] = 0$ for all $i \in N$.

2. Strong monotonicity clearly implies that the analogue of Young's *independence condition* holds: if $u\ B\ v$ and $v\ B\ u$ for i, that is, if $u_i = v_i$ and $\Delta(u) = \Delta(v)$, then $f_i[u] = f_i[v]$.

3. The *null agent property* holds for f: if i is a null agent in u then $f_i[u] = 0$. Indeed, since $u_i = 0_i$ and $\Delta(u) = 0 = \Delta(0)$, the independence condition applies with $v = 0$ and gives $f_i[u] = f_i[0] = 0$ by step 1.

4. Let us consider a new basis $\mathcal{B}_* = \{u_*^1, u_*^2, \ldots, u_*^n, u_*^N\}$ where

- $u_*^k = (0, 0, \ldots, 0, \overset{k}{\smile}{1}, 0, \ldots, 0|1)$ for $k = 1, 2, \ldots, n$ and
- $u_*^N = (0, 0, \ldots, 0|1)$.

If $u = (u_1, u_2, \ldots, u_n|u_N)$ then its expression in this basis is

$$u = u_1 u_*^1 + u_2 u_*^2 + \ldots, u_n u_*^n + \Delta(u) u_*^N. \tag{7}$$

5. The action of f on a multiple of any member of $\mathcal{B}_*$ is determined: for any $\lambda \in \mathbb{R}$,

- $f[\lambda u_*^k] = (0, 0, \ldots, 0, \overset{k}{\smile}{\lambda}, 0, \ldots, 0)$ for $k = 1, 2, \ldots, n$ and
- $f[\lambda u_*^N] = \lambda n(1, 1, \ldots, 1)$,

having used efficiency in both cases, the null agent property (step 3) in the former and symmetry in the latter.

6. Finally, we shall prove that f is determined on any $u = (u_1, u_2, \ldots, u_n|u_N) \in E_{n+1}$. Let $I = I(u)$ be the number of nonzero terms that appear in Eq. (7).

$I(u) = 0$: then $u = 0$ and step 1 applies.

$I(u) = 1$: then u is a nonzero multiple of a member of $\mathcal{B}_*$ and step 5 applies.

Let $I(u) \geq 2$. Therefore two cases arise:

$$\text{(a)} \quad u = \sum_{k \in K} u_k u_*^k \quad \text{and} \quad \text{(b)} \quad u = \sum_{k \in K} u_k u_*^k + \Delta(u) u_*^N,$$

for some nonempty $K \subseteq N$ such that $u_k \neq 0$ for all $k \in K$ in both cases, with $|K| \geq 2$ in case (a) and $|K| \geq 1$ and $\Delta(u) \neq 0$ in case (b).

Case (a). Here $u_j = 0$ for all $j \notin K$, and $\Delta(u) = 0$ because u is additive. Therefore, each $j \notin K$ is a null agent in u and hence $f_j[u] = 0$ by step 3. Now, if $i \in K$, let us take $v = u_i u_*^i$. Then $v_i = u_i$ and $\Delta(v) = 0 = \Delta(u)$ because v is additive. Thus $u\,B\,v$ and $v\,B\,u$ for i, so that, by the independence condition (step 2), $f_i[u] = f_i[v]$, which is determined by step 5.

Case (b). Assume, first, that $j \notin K$. Then $u_j = 0$ and let us take $v = \Delta(u) u_*^N$. Thus $v_j = 0 = u_j$ and $\Delta(v) = \Delta(u)$, and hence $u\,B\,v$ and $v\,B\,u$ for j. By the independence condition, $f_j[u] = f_j[v]$, which is determined by step 5. If, moreover, $I(u) = 2$, then $K = \{i\}$ for some $i \in N$ and $f_i[u]$ is determined by efficiency. If, instead, $I(u) > 2$, let $i \in K$ and $v = u_i u_*^i + \Delta(u) u_*^N$. Then $v_i = u_i$ and $\Delta(v) = \Delta(u)$, so that $u\,B\,v$ and $v\,B\,u$ for i and the independence condition yields $f_i[u] = f_i[v]$, which is already determined because v is of the form (b) and $I(v) = 2$. □

Remark 5.5 (a) A simple counterexample shows that the proportional rule does not satisfy strong monotonicity: if $n = 2$, $u = (9, 10|29)$ and $v = (8, 1|19)$ then $u\,B\,v$ for agent 1 but $\pi_1[u] = 13.7368 < 16.8889 = \pi_1[v]$.

(b) The members of basis $\mathcal{B}_*$ used in the above proof are in correspondence with the unanimity games that span $\mathrm{Im}(\sigma)$ as mentioned in Remark 4.2: the dictatorships $v_{\{1\}}, v_{\{2\}}, \ldots, v_{\{n\}}$, which are additive but not symmetric games, and the full unanimity game v_N, which is a symmetric and superadditive game. Eq. (7) shows that $\mathcal{B}_*$ is the most natural basis for E_{n+1}.

6 Criticism on the Proportional Rule

We shall discuss here several aspects of the proportional rule, most of which are far from being satisfactory, and will contrast them with the behavior of the Shapley rule.

6.1 Restricted Domain

As was already mentioned in Definition 2.2, the domain of the proportional rule π is not the entire space E_{n+1} but the subset defined by Eq. (3):

$$E^{\pi}_{n+1} = \{u \in E_{n+1} : u_1 + u_2 + \cdots + u_n \neq 0\}.$$

For instance, in Example 2.1 π applies to u^c and gives the (rounded) sharing of the total cost of -966,

$$\pi_1[u] = -176.28, \quad \pi_2[u] = -394.86 \quad \text{and} \quad \pi_3[u] = -394.86,$$

but it cannot be applied to u^s.

By contrast, the Shapley rule $\overline{\varphi}$ applies to all PBPs without restriction. In the case of Example 2.1 it gives

$$\overline{\varphi}_1[u^c] = -156.67, \quad \overline{\varphi}_2[u^c] = -404.67 \quad \text{and} \quad \overline{\varphi}_3[u^c] = -404.67$$

for the cost PBP u^c and

$$\overline{\varphi}_1[u^s] = 43.33, \quad \overline{\varphi}_2[u^s] = 43.33 \quad \text{and} \quad \overline{\varphi}_3[u^s] = 43.33$$

for the saving PBP u^s. This reflects the fairness of the Shapley rule: equal sharing of savings. We will see below that this fairness agrees with the consistency of the Shapley rule.

6.2 Doubly Discriminatory Level

Within its domain, the proportional rule coincides with the Shapley rule just on additive or symmetric PBPs. However, these are very particular cases and, in general, the two rules differ. As a matter of comparison, note that the expression of $\pi_i[u]$ given in Eq. (2),

$$\pi_i[u] = u_i + \frac{u_i}{\sum_{j \in N} u_j} \Delta(u),$$

shows that the proportional rule (a) allocates to each agent his individual utility (as the Shapley rule does) but (b) it shares the remaining utility proportionally to the individual utilities. In other words, no solidarity component exists in the proportional rule, as both components are of a competitive nature. Instead, in this second step the Shapley rule acts equitably (notice that the calculus for the Shapley rule is, therefore, easier than for the proportional rule). Then the proportional rule is, conceptually, more complicated than the Shapley rule and may include a *doubly discriminatory level* since, when comparing any two agents, it rewards twice the agent that individually can get the highest utility on his own. This discriminatory level arises, for example, in the case of nonnegative and superadditive PBPs in E^{π}_{n+1}. It is hard to find a reasonable justification for this.

6.3 The Axiomatic Framework

On its restricted domain E^{π}_{n+1}, where the Shapley rule has been axiomatically characterized by Theorem 5.3, the proportional rule satisfies the properties of efficiency, null and dummy agent and symmetry. It fails to satisfy additivity (otherwise, it would coincide with the Shapley rule by Theorem 5.3) and also strong monotonicity.

Now, in spite of its simplicity and mathematical tradition, it may be that additivity is, in principle, the least appealing property and might seem to practitioners only a “mathematical delicatessen”: the reason is that one does not easily capture the meaning of the sum of PBPs in practice. This will be illustrated in the next subsections. Incidentally, notice that π coincides with $\overline{\varphi}$ on each member of the basis $\mathcal{B}_0$ but this coincidence cannot be extended to all PBPs precisely because π is not additive. And, while the Shapley rule is linear, the proportional rule is only *homogeneous*, i.e. satisfies $\pi[\lambda u] = \lambda \pi[u]$ for every real number $\lambda \neq 0$ and every $u \in E^{\pi}_{n+1}$.

Although individual rationality has not been used in our axiomatic systems, it could be added here that the Shapley rule satisfies this property for all additive or superadditive (the most interesting) PBPs, whereas the proportional rule does not.

Table 1 Purchasing pool data

	Firm	Order cost u^0	Discounts applied (%)	Actual cost u^c	Saving u^s
	{1}	−1500	9	−1455	45
	{2}	−2400	9 and 15	−2250	150
	{3}	−3000	9 and 15	−2760	240
Pool	{1,2,3}	−6900	9, 15, and 24	−5724	1176

6.4 Inconsistency: Cost-Saving Problems

In Example 2.1, where related costs and savings arise, the proportional rule cannot be applied to the saving PBP and no kind of consistency can then be discussed. Instead, the consistency of the Shapley rule is clear since, for each $i \in N$,

$$\overline{\varphi}_i[u^c] = u_i + \overline{\varphi}_i[u^s].$$

The conclusion is that, using the Shapley rule, all towns are indifferent between sharing costs and sharing savings (as it should be).

Example 6.1 (*A purchasing pool*) Here the proportional rule will apply to all PBPs, but it will show inconsistency. Let $N = \{1, 2, 3\}$ be a purchasing pool of three firms and assume that, periodically, its members make to a common supplier orders of 1500, 2400 and 3000 units, respectively, of a product with unit cost 1. The supplier offers the following discounts:

- nothing for units from 1 to 1000
- 9 % off for units from 1001 to 2000
- 15 % off for units from 2001 to 3000
- 24 % off for units from 3001 upwards

Table 1 provides the full data for this purchasing pool. The members of the pool do not form a joint venture. They join just to get discounts for accumulated orders. Two alternatives are offered: (a) sharing the actual joint cost of −5724; (b) sharing the joint saving of 1176 after assuming that, previously, all members have individually deposited in a joint bank account the cost of their respective orders without discounts and the supplier's bill has been already paid from this account.

Table 2 provides the result of applying the proportional and Shapley rules to each alternative. Notice that we have three PBPs: an additive PBP u^0 of costs without discount, a PBP u^c of actual costs (i.e., with discount), and a PBP u^s of savings. They are obviously related by $u^0 + u^s = u^c$. While the Shapley rule is consistent in the sense that $\overline{\varphi}[u^0] + \overline{\varphi}[u^s] = \overline{\varphi}[u^c]$, which follows from additivity, this is not the case for the proportional rule, which does not satisfy this property as can be checked in Table 2. Notice, moreover, that $\pi[u^0] = \overline{\varphi}[u^0]$ because u^0 is an additive PBP.

Therefore, when using the Shapley value all members of the pool are indifferent between sharing costs with discount and sharing savings. Instead, this is not the case if the proportional rule is applied: firms 1 and 2 prefer sharing costs whereas

Table 2 Purchasing pool allocations

i	$\pi_i[u^0]$	$\pi_i[u^c]$	$\pi_i[u^s]$	$\overline{\varphi}_i[u^0]$	$\overline{\varphi}_i[u^c]$	$\overline{\varphi}_i[u^s]$
1	−1500	−1288.23	121.66	−1500	−1208	292
2	−2400	−1992.11	405.52	−2400	−2003	397
3	−3000	−2443.66	648.83	−3000	−2513	487
Sums	−6900	−5724.00	1176.00	−6900	−5724	1176

Table 3 Water and gas supply

Group	Water costs u^w	Gas costs u^g	Water + gas added costs $u^w + u^g$
{1}	−300	−150	−450
{2}	−200	−500	−700
{3}	−100	−250	−350
{1,2,3}	−540	−720	−1260
{1}+{2}+{3}	−600	−900	−1500

Table 4 Water and gas allocations

i	$\pi_i[u^w]$	$\pi_i[u^g]$	$\pi_i[u^w + u^g]$	$\overline{\varphi}_i[u^w]$	$\overline{\varphi}_i[u^g]$	$\overline{\varphi}_i[u^w + u^g]$
1	−270	−120	−378	−280	−90	−370
2	−180	−400	−588	−180	−440	−620
3	−90	−200	−294	−80	−190	−270
Sums	−540	−720	−1260	−540	−720	−1260

firm 3 prefers sharing savings; the inconsistency (or lack of fairness) of the procedure is obvious.

6.5 Inconsistency: Added Costs Problems

Let us consider a second example where additivity plays a crucial role.

Example 6.2 (*Added costs*) We slightly modify Example 2.1 and assume that the towns are interested in two goods (say, water and gas) carried by the same supplier: the costs are given in Table 3.

Again, we consider three PBPs: u^w, which describes the water costs; u^g, which gives the gas costs; and the sum $u^w + u^g$ that yields the added costs. Table 4 provides the result of applying the proportional and Shapley rules to each one of these PBPs. While the Shapley rule is consistent in the sense that $\overline{\varphi}[u^w + u^g] = \overline{\varphi}[u^w] + \overline{\varphi}[u^g]$, as follows from additivity, this is not the case for the proportional rule, which does not satisfy this property and fails therefore to be consistent in added costs problems.

In this case, if the proportional rule is applied town 1 prefers to share the payment of a water + gas joint bill, whereas towns 2 and 3 prefer to share the payment of separate bills. Instead, using the Shapley rule all towns are indifferent between sharing separate bills and sharing a joint bill.

7 Conclusions

The axiomatic viewpoints established by Shapley when defining the value notion for cooperative games, and by Young when replacing the dummy/null player and additivity properties with strong monotonicity, which have been adapted to PBPs, allow us to evaluate any sharing rule and, in particular, to compare the proportional rule and the Shapley rule. The relevant points are the following:

1. A first essential failure of the proportional rule is its restricted domain, defined by Eq. (3). Instead, the Shapley rule applies without any restriction to all PBPs.
2. When putting together Eqs. (2) and (6), the procedures look somewhat similar: first, each agent i is allocated his individual utility u_i; then, the surplus is shared among all agents. However, it is worth mentioning that the Shapley rule shares the surplus in equal parts, whereas the proportional rule shares it in proportion to the individual utilities. This means that the proportional rule is, conceptually, more complicated than the Shapley rule and includes a doubly discriminatory level that rewards twice the agent that individually can get the highest utility by his own. We cannot find a reasonable justification for this.
3. In which cases do these two allocation rules coincide? As has been shown, the Shapley rule and the proportional rule coincide on a PBP u (satisfying Eq. (3), of course) iff this PBP is additive or symmetric—the most trivial cases.
4. As to the Shapley axioms, in its restricted domain defined by Eq. (3) the proportional rule satisfies the properties of efficiency, dummy and null agent, and symmetry. It does not satisfy the strong monotonicity property in Young's sense.
5. This leaves us with the lack of additivity for this rule (otherwise, it would coincide with the Shapley rule by uniqueness on E^{π}_{n+1} according to Theorem 5.3). Thus the proportional rule is homogeneous but not linear. Let us raise the following question: is this failure important or, on the contrary, is additivity simply a standard mathematical property, just of a technical nature, without special relevance for practitioners? The answer is quite surprising. From the lack of additivity, serious inconsistencies of the proportional rule follow when applying it to certain problems. Examples 6.1 and 6.2 have illustrated this.

In summary, we have analyzed the proportional rule, from an axiomatic viewpoint but also from a practical viewpoint. Several properties and failures of the proportional rule have been remarked and, especially, practical implications of the non-additivity of this rule have been evidenced that result in a serious inconsistency when dealing with e.g. related costs-savings problems and added costs problems.

We therefore contend that the Shapley rule should replace in practice the proportional rule in PBPs, that is, in cooperative affairs where the coalitions of intermediate size ($1 < |S| < n$) do not matter.

The advantages of the Shapley value or rule over the proportional rule or value are even greater when considering general cooperative games and not only quasi-additive games (i.e. PBPs). The proportional rule does not take into account most of the coalitional utilities: precisely, all those corresponding to the intermediate coalitions. This becomes more and more critical as the number of players increases, and it gives rise to a very low sensitivity. On the contrary, the Shapley value is always concerned with all marginal contributions without exception and enjoys therefore a nice sensitivity with regard to the data defining any given problem.

Acknowledgments Research partially supported by Grants SGR 2009–01029 of the Catalonia Government (*Generalitat de Catalunya*), and MTM 2012–34426 of the Economy and Competitiveness Spanish Ministry.

References

Aadland, D., & Kolpin, V. (1998). Shared irrigation costs: An empirical and axiomatic analysis. *Mathematical Social Sciences*, *849*, 203–218.
Bergantiños, G., & Vidal-Puga, J. J. (2007). The optimistic TU game in minimum cost spanning tree problems. *International Journal of Game Theory*, *36*, 223–239.
van den Brink, R., & Funaki, Y. (2009). Axiomatizations of a class of equal surplus sharing solutions for TU-games. *Theory and Decision*, *67*, 303–340.
Carreras, F., & Freixas, J. (2000). A note on regular semivalues. *International Game Theory Review*, *2*, 345–352.
Driessen, T. S. H., & Funaki, Y. (1991). Coincidence of and collinearity between game theoretic solutions. *OR Spektrum*, *13*, 15–30.
Gillies, D. B. (1953). *Some theorems on n-person games*. Ph.D. Thesis. Princeton University.
Hart, S., & Mas-Colell, A. (1988). The potential of the shapley value. In A. E. Roth (Ed.), *The shapley value: essays in honor of Lloyd S. Shapley* (pp. 127–137). Cambridge: Cambridge University Press.
Hart, S., & Mas-Colell, A. (1989). Potential, value and consistency. *Econometrica*, *57*, 589–614.
Isbell, J. R. (1956). A class of majority games. *Quarterly Journal of Mathematics*, *7*, 183–187.
Kalai, E., & Smorodinsky, M. (1975). Alternative solutions to Nash's bargaining problem. *Econometrica*, *43*, 513–518.
Kalai, E. (1977). Proportional solutions to bargaining situations: Interpersonal utility comparisons. *Econometrica*, *45*, 1623–1630.
Megiddo, N. (1974). On the nonmonotonicity of the bargaining set, the kernel, and the nucleolus of a game. *SIAM Journal of Applied Mathematics*, *27*, 355–358.
Nash, J. F. (1950). The bargaining problem. *Econometrica*, *18*, 155–162.
von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton: Princeton University Press.
Ortmann, K. M. (2000). The proportional value for positive cooperative games. *Mathematical Methods of Operations Research*, *51*, 235–248.
Owen, G. (1995). *Game theory*. New York : Academic Press.
Roth, A. E. (Ed.). (1988). *The shapley vlue: Essays in honor of Lloyd S*. Shapley: Cambridge University Press.
Shapley, L. S. (1953). A value for *n*-person games. *Annals of Mathematical Studies*, *28*, 307–317.
Shapley, L. S. (1971). Cores of convex games. *International Journal of Game Theory*, *1*, 11–26.
Shubik, M. (1962). Incentives, decentralized control, the assignment of joint costs, and internal pricing. *Management Science*, *8*, 325–343.
Young, H. P. (1985). Monotonic solutions of cooperative games. *International Journal of Game Theory*, *14*, 65–72.

Veto Players and Non-Cooperative Foundations of Power in Legislative Bargaining

Andreas Nohn

In legislative bargaining of the Baron-Ferejohn type, veto players either hold all of the overall power of 1 and share proportional to their recognition probabilities, or hold no power at all. Hence, in this setting, it is impossible to provide non-cooperative support for power indices that do not assign all or no power to veto players. This highlights problems in the interpretation of results of Laruelle and Valenciano (2008a, b) which are taken as support for the Shapley-Shubik index and other normalized semi-values.

1 Introduction

An important and often raised question for cooperative solution concepts such as power indices is whether they can be given non-cooperative support or, equivalently, non-cooperative foundations. By this it is meant the existence of a non-cooperative game which, firstly, resembles the cooperative situation at hand in a reasonable way and, secondly, possesses equilibria whose payoffs coincide with those selected by the solution concept. As for power indices, the support of the Shapley-Shubik index (Shapley and Shubik 1954) has received the greatest attention. Foundations for the more general Shapley value (Shapley 1953), of which the Shapley-Shubik index is the restriction to simple games, have been

This chapter has been published in Homo Oeconomicus 29 (2), 2012.

A. Nohn
Public Choice Research Centre, Turku, Finland

A. Nohn (✉)
Institute of Socio Economics, University of Hamburg, Hamburg, Germany
e-mail: nohn@econ.uni-hamburg.de

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_37, © Springer-Verlag Berlin Heidelberg 2013

provided by Hart and Mas-Colell (1996) and Pérez-Castrillo and Wettstein (2001), among others.[1]

Legislative bargaining as introduced by Baron and Ferejohn (1989) and later extended by Eraslan (2002) and Eraslan and McLennan (2011) considers in its non-cooperative description of simple games only the most essential features of bargaining. For this reason, it is of particular interest whether power indices can be supported in this setting. Based on a result for the power of veto players, this note gives a partial, negative answer to this question. In addition, it highlights problems in the interpretation of results of Laruelle and Valencian (2008a, b, Sect. 4.4) which are taken as support for the Shapley-Shubik index and other normalized semi-values.

The note is organized as follows. Section 2 introduces legislative bargaining and existent results. The power of veto players is discussed in Sect. 3. Section 4 defines suitable notions and derives a negative result for the non-cooperative support of power indices in legislative bargaining.

2 Legislative Bargaining

This section presents legislative bargaining games and existing results. Notations and formulations mostly resemble Eraslan (2002) or Montero (2006), yet only Eraslan and McLennan (2011) establish the results in this generality.

A *legislative bargaining game* is given by a tuple (N, W, p, δ) where (N, W) is a *simple game*, i.e. N is the non-empty and finite set of *players* and $W \subseteq 2^N$ the set of *winning coalitions* for which $\emptyset \notin W$, $S \in W$ implies $T \in W$ for all $T \supseteq S$, and $N \in W$.[2] The *protocol* $p = (p_i)_{i \in N}$ on N with $p_i \geq 0$ for all $i \in N$ and $\sum_{i \in N} p_i = 1$ defines the *recognition probabilities* of players, and δ with $0 \leq \delta < 1$ is the *discount factor* common to all players. Then, in every of possibly infinitely many rounds, one player $i \in N$ is randomly chosen as the *proposer* according to protocol p. Proposer i chooses a feasible coalition $S \in W_i = \{S \in W | S \ni i\}$ and suggests a split of a unit surplus between the members of S. In some arbitrary order, all *respondents* $j \in S \backslash \{i\}$ vote on acceptance or rejection of this offer. Unless all respondents agree, the game proceeds to the next round and future payoffs are discounted by δ. If respondents agree unanimously, the suggested split is implemented and the game ends. In addition, players are assumed to be risk-neutral, i.e. payoffs equal the (possibly discounted) shares received, and to have complete and perfect information.

---

[1] Gul (1989) provides another foundation for the Shapley value. However, his assumptions do not apply to simple games, and hence his results cannot be interpreted as support for the Shapley-Shubik index.

[2] It is *not* necessary for the results of this note to require *properness* of the simple game, i.e., that $S \in W$ implies $N \backslash S \notin W$.

Now, restrict attention to *stationary strategies* for which actions do not depend on past rounds. Thus, equilibrium strategies can be characterized by the corresponding (and as well stationary) *payoffs* $v = (v_i)_{i \in N}$ that players expect before any round. Facing the choice of a coalition, proposer i chooses a coalition according to his *selection distribution* $\lambda_i = (\lambda_{iS})_{S \in W_i}$ with positive probability only on feasible coalitions $S \in W_i$ which have maximum *excess* $1 - \sum_{j \in S} \delta v_j$. Having chosen coalition S, proposer i offers any respondent $j \in S \setminus \{i\}$ that player's *continuation value* δv_j, his own share amounting to his own continuation value δv_i and the coalition's excess $1 - \sum_{j \in S} \delta v_j$. Irrespective of proposer or coalition, a respondent j accepts any share equal to or greater than his continuation value δv_j.

Thus, payoffs v are supported by a *stationary subgame perfect equilibrium* (SSPE) with selection distributions $(\lambda_i)_{i \in N}$ if and only if, for all $i \in N$,

$$v_i = p_i \sum_{S \in W_i} \lambda_{iS}(1 - \sum_{j \in S} \delta v_j) + \sum_{j \in N} p_j \sum_{S \in W_i \cap W_j} \lambda_{jS} \delta v_i \tag{1}$$

and $\lambda_i = (\lambda_{iS})_{S \in W_i}$ puts positive probability only on feasible coalitions $S \in W_i$ that have maximum excess $1 - \sum_{j \in S} \delta v_j$.[3] Eraslan and McLennan (2011) show the existence of a SSPE and the uniqueness of corresponding payoffs. As follows from above characterization of stationary equilibrium strategies, there is no delay of agreement. The SSPE payoffs are nonnegative and efficient, $v_i \geq 0$ for all $i \in N$ and $\sum_{i \in N} v_i = 1$.

3 The Power of Veto Players

Given simple game (N, W) and protocol p on N, denote $\xi(N, W, p) = (\xi_i(N, W, p))_{i \in N}$ as the limit of SSPE payoffs in legislative bargaining games (N, W, p, δ) as $\delta \to 1$.[4] Commonly, $\xi_i(N, W, p)$ is interpreted as the *(bargaining) power* of player $i \in N$. Power $\xi(N, W, p)$ is nonnegative and efficient, $\xi_i(N, W, p) \geq 0$ for all $i \in N$ and $\sum_{i \in N} \xi_i(N, W, p) = 1$. A *veto player* $i \in N$ is a member of all winning coalitions, $i \in S$ for all $S \in W$. Denote the set of veto players by V.

Proposition 1 *Let* (N, W) *be a simple game and* p *a protocol on* N. *Assume the set of veto players* V *is non-empty and* $p_i > 0$ *for some* $i \in V$. *Then veto players hold all power and share it proportional to their recognition probabilities,*

---

[3] In Eq. (1), player i's share from being proposer is split into excess $1 - \sum_{j \in S} \delta v_j$ in the first sum and his continuation value δv_i in the second sum. This presentation of the formula is more convenient for the proof of proposition 1.

[4] The theorem of the maximum ensures that the unique payoffs are continuous in δ. Hence, the limit of payoffs as $\delta \to 1$ exists.

$\xi_i(N,W,p) = p_i / \sum_{j \in V} p_j$ for all $i \in V$. In particular, all other players have no power, $\xi_i(N,W,p) = 0$ for all $i \notin V$.[5]

Proof The proposition is shown in two steps, (i) $p_j \xi_i(N,W,p) = p_i \xi_j(N,W,p)$ for all $i,j \in V$ and then (ii) $\sum_{i \in V} \xi_i(N,W,p) = 1$. The two statements combined yield $\xi_i(N,W,p) = p_i / \sum_{j \in V} p_j$ for all $i \in V$.

(i) For an arbitrary discount factor δ, consider legislative bargaining game (N,W,p,δ) with SSPE payoffs v and selection distributions $(\lambda_i)_{i \in N}$. For a veto player $i \in V$, $\sum_{j \in N} p_j \sum_{S \in W_i \cap W_j} \lambda_{jS} = 1$ and the optimality of i's selection distribution λ_i allow to write (1) as $v_i = p_i \max_{S \in W}(1 - \sum_{j \in S} \delta v_j)/(1-\delta)$. So, since v_i depends on $i \in V$ only through p_i, it is $p_j v_i = p_i v_j$ for all $i,j \in V$. Hence, as $\delta \to 1$, $p_j \xi_i(N,W,p) = p_i \xi_j(N,W,p)$ for all $i,j \in V$.

(ii) Fix $i \in V$ with $p_i > 0$. Without loss of generality, assume the selection distributions of players are continuous in δ and denote $\lambda_j^* = (\lambda_{jS}^*)_{S \in W}$ as the limit of selection distributions of any player $j \in N$ as $\delta \to 1$.[6] The limit of i's equation system (1) for $\delta \to 1$ then reads

$$\xi_i(N,W,p) = p_i \sum_{S \in W_i} \lambda_{iS}^* (1 - \sum_{j \in S} \xi_j(N,W,p)) + \sum_{j \in N} p_j \sum_{S \in W_i \cap W_j} \lambda_{jS}^* \xi_i(N,W,p).$$

Noticing $\sum_{j \in N} p_j \sum_{S \in W_i \cap W_j} \lambda_{jS}^* = 1$ yields $\sum_{S \in W_i} \lambda_{iS}^* (1 - \sum_{j \in S} \xi_j (N,W,p)) = 0$. Given this, $W_i = W$ and the optimality of λ_i^* requires $\sum_{j \in S} \xi_j(N,W,p) = 1$ for all $S \in W$. Now, for all $j \notin V$, there is $S \in W$ with $S \not\ni j$ and thus $\xi_j(N,W,p) = 0$. In particular, $\sum_{j \in V} \xi_j(N,W,p) = 1$. □

Infinitely patient and able to block all proposals, veto players can play off the remaining players against each other to an extent that none of the latter can expect any positive share from bargaining. To do so, however, veto players themselves need an at least arbitrarily small possibility to make offers. In the case of V being non-empty but $p_i = 0$ for all $i \in V$, the power $\xi_i(N,W,p)$ of any veto player $i \in V$ is 0.[7]

[5] Note that, in footnote 9 of her chapter, Montero (2006) also claims, however does not proof, that veto players hold all power.

[6] The theorem of the maximum allows to choose selection distributions such that they are continuous in δ and such that their limit exists.

[7] In any legislative bargaining game (N,W,p,δ), the payoff of any player $i \in N$ with $p_i = 0$ is zero, $v_i = 0$ (see for instance Eraslan and McLennan 2011). This of course translates to i's power for simple game (N,W) and protocol p.

4 On the Support of Power Indices

Denote a *protocol scheme* p as a mapping which assigns to every simple game (N, W) a protocol $p(N, W) = (p_i(N, W))_{i \in N}$ on N. Thus, $\xi(N, W, p(N, W))$ is the bargaining power supported by protocol scheme p for simple game (N, W). A *power index* f is a mapping which assigns a distribution of power $f(N, W) = (f_i(N, W))_{i \in N}$ to every simple game (N, W). Then, protocol scheme p *supports* power index f if $\xi(N, W, p(N, W)) = f(N, W)$ for every simple game (N, W). Since power is nonnegative and efficient, only nonnegative and normalized power indices can be supported by a protocol.

Given some power index that does not assign all or no power to veto players, proposition 1 and the subsequent comment yield that any protocol scheme fails to support such a power index—at least in simple games with veto players, induced power and power index do not coincide.

Corollary 1 There exists no protocol scheme supporting any power index that does not assign all or no power to veto players.

In particular, legislative bargaining as considered here does not allow for support of the Shapley-Shubik index or the (normalized) Banzhaf index (Banzhaf 1965), Deegan-Packel index (Deegan and Packel 1978), or Public Good index (Holler 1982). However, there is the possibility that other solution concepts which do ascribe all power to veto players can be supported by suitable protocol schemes. For instance, Montero (2006) shows the nucleolus (Schmeidler 1969) is even *self-confirming*, i.e. it is supported by the protocol scheme which assigns the nucleolus itself as the protocol.

Corollary 1 also highlights a misinterpretation of results of Laruelle and Valenciano (2008a, b, Sect. 4.4). They investigate non-cooperative bargaining as in this note, apart from two exceptions. Firstly, they more generally allow for non-transferable utility, and secondly, they deal with unanimity bargaining only. In this setting, they consider the Shapley-Shubik index (and other normalized semi-values) of an arbitrary simple game as the protocol and find it then also emerging as bargaining power.[8] At first sight, the result itself suggests to be taken as non-cooperative support. The authors seem to understand it this way as well and even say '[it] provides a noncooperative interpretation of any reasonable power index' ((Laruelle and Valenciano 2008a, p. 352)). Only note that, in the particular case of unanimity bargaining with transferable utility, power is always given by the recognition probabilities: proposition 1 implies $\xi(N, \{N\}, p) = p$ for all sets of players N and arbitrary protocols p on N. So any nonnegative and normalized power index, be it 'reasonable' or not, could hence be supported by using it as a

[8] In the more general case of non-transferable utility, bargaining power materializes as the weights of the asymmetric Nash-solution, the latter arising as the limit of payoffs when players grow infinitely patient. In the particular case of transferable utility, bargaining power then again corresponds with payoffs exactly as in this note.

protocol scheme. However, this approach does not seem adequate. The modeling of an arbitrary voting situation as unanimity bargaining typically bears an incongruence that, as the results of this note show, can matter significantly.

Acknowledgments I am grateful for financial support from the Academy of Finland, and also want to thank Manfred J. Holler, Stefan Napel, Maria Montero as well as two anonymous referees for very valuable comments.

References

Banzhaf, J. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review, 19*, 317–343.

Baron, D., & Ferejohn, J. (1989). Bargaining in legislatures. *American Political Science Review, 83*, 1181–1206.

Deegan, J., & Packel, E. (1978). A new index of power for simple n-person games. *International Journal of Game Theory, 7*, 113–123.

Eraslan, H. (2002). Uniqueness of stationary equilibrium payoffs in the Baron-Ferejohn model. *Journal of Economic Theory, 103*, 11–30.

Eraslan, H., & McLennan, A. (2011). *Uniqueness of stationary equilibrium payoffs in coalitional bargaining*. Working paper.

Gul, F. (1989). Bargaining foundations of Shapley value. *Econometrica, 57*(1), 81–95.

Hart, S., & Mas-Colell, A. (1996). Bargaining and value. *Econometrica, 64*(2), 357–380.

Holler, M. (1982). Forming coalitions and measuring voting power. *Political Studies, 30*, 262–271.

Laruelle, A., & Valenciano, F. (2008a). Noncooperative foundations of bargaining power in committees and the Shapley-Shubik index. *Games and Economic Behavior*, *63*, 341–353.

Laruelle, A., & Valenciano, F. (2008b). *Voting and collective decision making: Bargaining and power*. Cambridge: Cambridge University Press.

Montero, M. (2006). Noncooperative foundations of the nucleolus in majority games. *Games and Economic Behavior*, *54*, 380–397.

Pérez-Castrillo, D., & Wettstein, D. (2001). Bidding for the surplus: A non-cooperative approach to the Shapley value. *Journal of Economic Theory*, *100*, 274–294.

Schmeidler, D. (1969). The nucleolus of a characteristic function game. *SIAM Journal on Applied Mathematics*, *17*(6), 1163–1170.

Shapley, L. (1953). A value for *n*-person games. In H. Kuhn & A. Tucker (Eds.), *Contributions to the theory of games, II. Vol. 28 of Annals of Mathematical Studies* (pp. 307–317). Princeton: Princeton University Press.

Shapley, L., & Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review*, *48*, 787–792.

Distortion-Free Logrolling Mechanism

Hannu Vartiainen

1 Introduction

Parties are an integral part of a modern representative democracy as they aggregate information and simplify the collective bargaining process. And so is "logrolling" under which influence or bargaining power is traded among the parties prior to formal decision making procedure. Is logrolling desirable from the viewpoint of the society? Among the most fundamental results in economics is that trading is an efficient way to allocate scarce economic resources. This is especially true in large economies, where individual market sides do not have any market power. But it is not at all clear whether markets for bargaining power fits into this idealization.

There are two primary concerns. First, since "bargaining power" is not an economic good, i.e. private and separable, one cannot analyze the agents' choices as if they are independent. As a consequence, the choice sets of the agents may no longer be convex, violating the fundamental prerequisites of the welfare theorems. Second, since the number of parties is usually small, the market for bargaining power fits badly to the idea of large markets. That is, it is unavoidable that a single party is left with market power which leaves the collective decision process vulnerable to exploitation and distortion.

In this chapter, we analyze the effectivity of logrolling in a particular set up, where the agents may trade bargaining power prior to the actual bargaining game. From the normative point of view, an important question is what is the right benchmark. One good reference point is the outcome that one would obtain in the absence of any party system, i.e., under *direct* bargaining when the members of the

I thank an anonymous referee for very useful comments. I have also benefitted from conversations with Klaus Kultti and Hannu Salonen.

H. Vartiainen (✉)
HECER, University of Helsinki, Arkadiankatu 7, 00014 Helsinki, Finland
e-mail: hannu.vartiainen@helsinki.fi

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_38,

society themselves distribute the resources, without parties.[1] At least two intuitive arguments suggest that direct bargaining should be *distortion-free*. First, when the society consists of many agents, the bargaining power of a single agent vanishes and the set up approaches the "perfect competition" ideal. Second, under direct decision making one does not need to worry about the incentive problems related to delegation. These points suggest that the theoretical outcome of direct bargaining serves as a good reference point for evaluating the goodness of bargaining through parties.

We study party formation and logrolling in a model where the society distributes its common resources, the "pie", via the Rubinstein unanimity bargaining game.[2] We compare two situations: (i) In the benchmark model, a large society distributes the resources via direct bargaining, i.e. each member of the society serves a player in the bargaining game. (ii) In the party formation model, we model the party system as a small (finite) number of groups that represent individuals with similar preferences (a party is composed of a group of identical individuals). The role of the parties is to bargain over common resources, and channel the benefits to their members.

First we establish formally the above conjecture: in the benchmark model (i), where the number of the agents is large but consists of a finite number of different "types", the relative importance of the first mover advantage vanishes and the bargaining outcome converges to a well defined *distortion-free allocation*. Conversely, when the number of bargainers is small, the distortion due to the first mover advantage cannot be avoided. Hence, we conclude that, in a society that consists of a large number of agents, delegating the bargaining right to a small number of parties creates a distortion.

Our main contribution is to develop a way to implement the distortion-free allocation even when the distribution of common resources is conducted by a small number of parties. We construct a two-stage logrolling mechanism, where the parties (or their representatives) *make bids* for the right to be the first proposer in the bargaining game. The party that is willing to give up most of its individual resources will get the right to serve as the first proposer. The key assumption is that the winning "bid" is *added* to the pool of social resources over which the actual bargaining takes place. The main result of the chapter is that when appended with the logrolling game, bargaining-through-parties induces the distortion-free allocation—the same that obtains under direct bargaining. Hence the effective distortion from a representative democracy that functions through a party system may not be significant. The role of logrolling is to restore the distortion-free allocation.

[1] There is a volumenous literature on the differences between electoral systems (see e.g. Baron 1989, 1993; Morelli 2004) but less attention has been paid to the fundamental question of why representative democracy is a dominant governing form at the first place.

[2] We use the n-player version of the model where all responders announce their acceptance in a sequential order and the first rejecting player gets the right to make the proposal in the next round (e.g. Herrero 1985). Our focus is on the *stationary* equilibrium of the game.

This suggests that the virtues of the representative party system such as flexibility, simplicity, and operationalizability may also be the *net* benefits of the system.

There are several important omissions in the model, each providing an interesting avenue for further research. An important one is that the domain of social decisions is matched with a unidimensional set of a shareable good. This permits a convex utility domain and, *a fortiori*, defines the distortion-free allocation. From the viewpoint of political theory, a more fitting scenario would perhaps be spatial preferences over a platform. This would restrict the feasible utility set and non-trivial implications on the bargaining outcome (see. Duggan and Cho 2003; Herings and Predtetchinski 2008).

A second omission is that the collective choice is based on a rather restrictive voting principle, the unanimity rule. It is warranted to ask whether the results survive when a more natural voting structure such as the majority rule (Baron and Ferejohn 1989) is assumed. The question remains open. While it is known that the majority based bargaining model supports a unique (stationary) equilibrium (Eraslan 2002) what is not known—and needed for our results to remain valid—is whether the equilibrium converges as the population (and hence number of majority coalitions) becomes large.[3] However, the results by Kalandrakis (2006) suggest that the details of the voting rule may not be crucial for the convergence.

A third restriction is the assumption that process of choosing representative agents is trivialized by the assumption that similar agents form a coalition. A more fundamental approach would allow endogenous candidacy as in a seminal appear by Besley and Coate (1997). Endogenous candidacy and delegated representation would also raise the issue of agency. Incentive costs associated to legislative bargaining are discussed e.g. Cai (2002) and Cai and Cont (2004).

From the viewpoint of economic theory, our result can be interpreted as a version of the Core convergence: the outcome that is obtained in large bargaining markets with negligent bargaining power can be simulated in a small market with a "Walrasian" auctioneer. However, the methods are quite different. In particular, there are no coalitions in our model—the spirit of our model is fully noncooperative.

This chapter is related to Kultti and Vartiainen (2007a) who study convergence of bargaining outcomes in a related model of large population. The driving force behind the convergence in there as well as in here is that, as the number of players becomes very large, the bargaining power of an individual player vanishes to zero. An important observation is that convergence has different characteristics than when the bargaining power vanishes due to speeding up the bargaining process (Binmore et al. 1986). In particular, the convergence point under large population is not related to the Nash bargaining solution.[4]

---

[3] Of course, there is only little hope to obtain a well defined voting solution in a completely general social choice set up. For a thorough survey on problems associated with voting in general setups, see Nurmi (1999).

[4] However, see Thomson and Lensberg (1989).

First we define the set up and specify the bargaining game. Then we establish the feasible arbitrations schemes. Finally, the implementation result is proven. Omitted proofs appear in the appendix.

2 Set up

Agents and resources There is a society of agents, distributing common resources. There are $1, \ldots, n$ agents, each of them endowed with one unit of resources. If an agent enters the society, then his resources become part of the common pool of resources.

Time preferences of the agent i has the representation $u_i(x_i)\delta^t$, where $x_i \in \mathbb{R}_+$ is the agent's consumption. We assume that the publicly observable utility functions $u_1, \ldots, u_n$ are drawn independently from a *finite* set U, whose cardinality is also denoted by U. The probability of the occurrence of $u \in U$ is λ_u, a rational number. We assume that each $u \in U$ is monotonously nonnegative, increasing, concave, and continuously differentiable function, and that $\delta \in (0, 1)$.[5]

A unanimity bargaining game For later purposes, we discuss of the bargaining game in a more general level than the set up of this section requires. Let the size of shareable resources be $X > 0$, and let the group of agents be a finite set N (whose cardinality we also denote by N). The set of allocations is

$$S = \left\{ x \in \mathbb{R}_+^N : \sum\nolimits_{i \in N} x_i \leq X \right\}.$$

Given N and X, we define a unanimity bargaining game $\Gamma^N(X, i)$ as follows: At any stage $t = 0, 1, 2, \ldots,$

- Player $i(t) \in N$ makes an offer $x \in S$. Players $j \neq i(t)$ accept or reject the offer in the ascending order of their index.[6]
- If all $j \neq i(t)$ accept, then x is implemented. If j is the first who rejects, then j becomes $i(t + 1)$.
- $i(0) = i$.

We focus on the *stationary subgame perfect equilibria*—simply equilibria or SPE in the sequel—of the game, where:

1. Each $i \in N$ makes the same proposal $x(i)$ whenever he proposes.
2. Each i's acceptance decision in period t depends only on x_i that is offered to him in that period.

---

[5] Weaker conditions would suffice (see Fishburn and Rubinstein 1982; or Kultti and Vartiainen 2007a, b). The current choice is for simplicity.

[6] The order in which players response to a proposal does not affect the results.

We now characterize the equilibria.[7] We first state an important intermediate result.

Lemma 1 *For any* $Y > 0$ *and* $c \in \mathbb{R}^{\mathbb{N}}_{++}$, *there is a unique* $x \in \mathbb{R}^{\mathbb{N}}_{++}$ *and* $d > 0$ *such that*

$$u_i(c_i x_i) = u_i((x_i + d)c_i)\delta, \textit{ for all } i \in N,$$
$$\textstyle\sum_{j \in N} x_j = Y.$$

The interpretation of the lemma is that for any $(c_1, \ldots, c_n)$, there is a unique $(x_1, \ldots, x_n)$ and d that add up to Y, and that have the property that each i is indifferent between the consumption $c_i x_i$ today and $(x_i + d)c_i$ tomorrow.

Lemma 2 *The unique stationary equilibrium outcome of* $\Gamma^N(X, i)$ *can be written* $(x_i + d, x_{-i})$ *where* $x \in \mathbb{R}^{\mathbb{N}}_{++}$ *and* $d > 0$ *such that*

$$u_i(x_i) = u_i(x_i + d)\delta, \quad \textit{for all } i \in N, \tag{1}$$

$$\textstyle\sum_{j \in N} x_j = X - d. \tag{2}$$

In Lemma 2, x_i is i's share of the pie of size X when he is in the role of a responder, and $x_i + d$ is his share when he is in the role of a proposer. Choosing $Y = X - d$ and combining Lemmata 1 and 2 the following result is obtained.

Proposition 1 *A stationary equilibrium of* $\Gamma^N(X, i)$ *exists. Moreover, the equilibrium is unique.*

Thus in our $n-$player context, the pool of shareable resources is n and the game is represented by $\Gamma^{\{1,\ldots,n\}}(n, i)$.

Corollary 1 *A stationary equilibrium of* $\Gamma^{\{1,\ldots,n\}}(n, i)$ *exists for any* n *and for any* i. *Moreover, the equilibrium is unique.*

3 Large Population, Parties, and the Distortion Free Allocation

We now establish that when the number of agents becomes large, the outcome of the bargaining procedure converges to a well defined limit. To characterize this limit, let $y^* \in \mathbb{R}^{\mathbb{U}}$ and $d^* > 0$ satisfy

[7] See also Krishna and Serrano (1996).

$$u(y_u^*) = u\big(y_u^* + d^*\big)\delta, \text{ for all } u \in U, \tag{3}$$

$$\textstyle\sum_{u\in U}\lambda_u y_u^* = 1. \tag{4}$$

Since each λ_u is a rational number, Lemmata 1 and 2 imply that the desired y^* and d^* do exist.[8]

The importance of (3–4) is that through them we can describe the limit outcome of the bargaining procedure in *finitary* terms (recall that U is a finite set). Define a *distortion-free allocation* $x^* = (x_1^*, x_2^*, \ldots) \in \mathbb{R}^\infty$ by the following condition:

$$x_i^* = y_u^* \text{ if } u = u_i, \text{ for all } i = 1, 2, \ldots, \text{ for all } u \in U. \tag{5}$$

Now, by (3–4), y^* can be interpreted as the expected division of all players' resource units, proposed by an arbitrator under the constraint that any player can reject the proposal and become the proposer in the next round.

Proposition 2 *As $n \to \infty$, the stationary equilibrium outcome of game $\Gamma^{\{1,\ldots,n\}}(n,i)$ converges to the distortion-free allocation x^*.*

Consider now the situation where bargaining is conducted via *coalitions* of similar agents—*political parties*. For each $u \in U$ all the agents of type u constitute a party, that represent the members of the group in a bargaining contest. Let each group select one agent as the representative of the group who is entitled to bargain and trade on behalf of the whole group. Gains and losses of the group are divided equally among its members.

Now let π_u be the portion of agents belonging to party u. If the portion of the good that is relegated to player i belonging to party u is x_u, then party u must obtain the share $x_u\pi_u$ of each of the n pies of all the agents. Thus game with parties can be interpreted as $\Gamma^{\pi U}(1,u)$; the set of shareable resources is 1, the index set of the players is U, the utility function of a representative player is $u((\cdot)\pi_u^{-1})$, and the party that begins the game is u. By Lemma 2, the unique stationary equilibrium outcome of the game is characterized by $(\xi_u^* + \varepsilon^*, \xi_{-u}^*)$ where $\xi^* \in \mathbb{R}^{\mathbb{U}}_{++}$ and $\varepsilon^* > 0$ such that

$$\begin{aligned} &u\big(\xi_u^*\pi_u^{-1}\big) = u\big((\xi_u^* + \varepsilon^*)\pi_u^{-1}\big)\delta, \text{ for all } u \in U, \\ &\textstyle\sum_{u\in U}\xi_u^* = 1 - \varepsilon^*. \end{aligned}$$

By the law of large numbers, the share of the agents in the u-group is λ_u as n becomes large, i.e. $\lim_n \pi_u = \lambda_u$. Letting $z_u^* = \xi_u^*\lambda_u^{-1}$ and $d_u^* = \varepsilon^*\lambda_u^{-1}$ for all u, we have, in the limit,

[8] Letting m be the least common denominator of the elements in $\{\lambda_u\}_{u\in U}$, (3)–(4) can be interpreted as the bargaining solution of the problem $\Gamma^{\{1,\ldots,m\}}(m+d,i)$, where the number of players with preferences u is $\lambda_u m$. The existence and uniqueness of the solution now follow by Lemma 2.

$$u\left(z_u^*\right) = u\left(z_u^* + d_u^*\right)\delta, \text{ for all } u \in U, \tag{6}$$

$$\sum\nolimits_{u \in U} \lambda_u z_u^* = 1 - d_u^* \lambda_u. \tag{7}$$

Thus we conclude that is cannot be the case that $z_i^* = y_u^*$ if $u = u_i$, for all $i = 1, 2, \ldots$, for all $u \in U$. If it were, then, by (3) and (6), $d^* = d_u^* > 0$ for all u, which implies a conflict between (4) and (7). Thus we conclude the following:

Proposition 3 *As $n \to \infty$, the stationary equilibrium outcome of bargaining that is conducted through parties does not converge to the distortion-free allocation x^*.*

Hence, if parties are an integral part of the society, simply changing the societal bargaining in a way that parties replace individuals in the game is not satisfactory. It leads to a distortion due to the first mover advantage. The next section shows that logrolling is a way to circumvent this problem.

4 Logrolling Mechanism

We now construct a simple "logrolling" mechanism that, when n becomes large, implements the distortion free allocation x^* even when bargaining is conducted through parties. Consider a market where the right to be the first proposer in a bargaining game is sold after a bidding contest to one of the U groups (or their representatives). The right is sold to the group that makes the highest bid (break ties by using randomization). Once the price p is paid by the winner it is added to the pool of resources over which bargaining then takes places.

The bidding contest can be interpreted as a logrolling game where all the groups, "parties", bid for the right to be the leader in the bargaining game, "political process", that follows the bidding contest. Only one group can serve as the initial proposer and hence enjoy from the bargaining power that comes with it.

More formally, since the agents' utility functions are i.i.d, λ_u is the limit share of type u agents in the population as the population becomes large. Since all gains and losses of the group are divided equally among its members, if z_u is the u-group's relative share of the total shareable resources, an u-type agent's consumption is approximated by $\lambda_u^{-1} z_u$ as n becomes large. It is convenient to describe the u-group's agents utilities directly in terms of z_u. The utility function $\bar{u}$ of the representative of the u-group with respect to z_u is:

$$\bar{u}(z_u) = u\left(\lambda_u^{-1} z_u\right), \text{ for all } z_u \in [0, 1]. \tag{8}$$

Function $\bar{u}$ is convex and continuous since u is.

Denote the set of normalized utility functions by $\bar{U}$. The rules of the logrolling mechanism Γ^* are formally as follows: Players in the set $\bar{U}$ first cast their bids.

Given the normalized resources 1, if $i \in \bar{U}$ wins the bidding contest with bid p, then the bargaining game $\Gamma^{\bar{U}}(1+p:i)$, with i as the first proposer, is initiated.

The interpretation of the logrolling mechanism is that the parties have to fight for the right to be in the leading position in the actual bargaining game (for example, as the prime minister). The party who is willing to sacrifice most of its own resources for the benefit of the whole society will win the contest. Our claim is that this mechanism implements the distortion-free arbitration scheme.

First, let $z_j(X)$ be what a receiver j gets in the game $\Gamma^{\bar{U}}(X:i)$. By (1) and (2) and Proposition 1, there is $z(X) = (z_1(X), \ldots, z_n(X))$ that is the unique solution to

$$\bar{u}(z_i(X)) = \bar{u}_i\left(X - \sum\nolimits_{j \neq i} z_j(X)\right)\delta \text{ for all } i. \tag{9}$$

By the Implicit Function Theorem, $z_i(\cdot)$ is continuous.

Lemma 3 *$z_i(X)$ is strictly increasing in X, for all i.*

By (1) and (2) there is a unique $(z^*_{\bar{u}})_{\bar{u}\in\bar{U}}$ and $p^* > 0$ such that

$$\bar{u}(z^*_{\bar{u}}) = \bar{u}\left(z^*_{\bar{u}} + p^*\right)\delta, \text{ for all } \bar{u} \in \bar{U}, \tag{10}$$

$$\sum\nolimits_{\bar{u}\in\bar{U}} z^*_{\bar{u}} = 1. \tag{11}$$

Lemma 4 *In equilibrium of the logrolling mechanism Γ^*, p^* is the winning bid and $z^*_{\bar{u}}$ is the $u-$group's share of resources, for z^* and p^* as specified in (10) and (11).*

We now argue that from the viewpoint of a single agent, the outcome of the auction among the representatives is the same as the limit outcome of the direct democracy. Hence the logrolling mechanism implements the distortion-free allocation.

Proposition 4 *The distortion-free allocation x^* is the unique equilibrium allocation of the logrolling mechanism Γ^*.*

5 Concluding Remarks

In a large society direct bargaining over resources leads to an outcome where one agent's ability to influence the outcome diminishes. This is desirable since then also the agent's ability to exploit their bargaining power as a first mover vanishes. In this chapter, the resulting outcome is called *distortion-free*. In contrast, when collective choice large society is made manageable by allocating the resources through bargaining by parties that are formed of agents with similar preferences, distortions cannot be avoided. The identity of the first mover has a real effect on the final outcome.

The main result of this chapter is to show that a *logrolling* mechanism, under which parties bid for the right to be the first proposer in the bargaining game, implements the distortion-free allocation. Under the logrolling mechanism a party, or its representative, acts on behalf of its members—gains and losses of the party are divided evenly. Before parties engage into bargaining over common resources, they compete over the right to make the first proposal in the bargaining game. The winning bid is added to the pool of common resources that is later shared via bargaining. The equilibrium outcome of this process coincides with the distortion-free allocation.

Thus logrolling can be thought as a mechanism that helps to restore the desired distortion-free outcome in a large society where collective choice cannot be managed without giving some party distortative bargaining power. The connection of this result to the First Welfare Theorem is clear.

Appendix: Proofs

Proof of Lemma 3. Define a function $\bar{v}$ such that

$$\bar{u}(\bar{v}(z_u)) = \bar{u}(z_u)\delta, \text{ for all } z_u \in [0, 1].$$

By the concavity of $\bar{u}$, $\bar{v}$ and $x - \bar{v}(x)$ are both increasing functions. Rewrite condition (9) as

$$\bar{v}_i^{-1}(z_i(X)) - z_i(X) = X - \sum\nolimits_{i \in U} z_j(X).$$

Since $x - \bar{v}(x)$ is an increasing function, and z_i continuous, z_i is strictly increasing if $X - \sum z_j(X)$ is. Since this applies to all i, $\sum z_j(X)$ is strictly increasing if $X - \sum z_j(X)$ is. But then, since $\sum z_j(X)$ being weakly decreasing means that $X - \sum z_j(X)$ is strictly increasing, it cannot be the case that $\sum z_j(X)$ is not strictly increasing. Thus $\sum z_j(X)$ is strictly increasing and hence z_i is strictly increasing.■

Proof of Lemma 4 *Only if*: First we argue that there are at least two highest bids. Suppose that there is a single highest bid. Then buying the proposing right with price p must be at least profitable as the opportunity cost of lowering the bid by a small $\varepsilon > 0$:

$$\left[1 + p - \sum_{j \neq i} z_j(1 + p)\right] - p \geq \left[1 + p - \varepsilon - \sum_{j \neq i} z_j(1 + p - \varepsilon)\right] - (p - \varepsilon).$$

That is

$$0 \geq \sum\nolimits_{j \neq i} [z_j(1 + p) - z_j(1 + p - \varepsilon)].$$

But by Lemma 3 this cannot hold.

Thus at least two bidders bid the winning bid p. Then buying the proposing right under p must be at least profitable as the opportunity cost of letting the other highest bidder win with price p:

$$\left[1 + p - \sum_{j \neq i} z_j(1 + p)\right] - p \geq z_i(1 + p). \tag{12}$$

Since increasing ones bid is not profitable for the losing bargainer j that bids p,

$$\left[1 + p + \varepsilon - \sum_{k \neq j} z_k(1 + p + \varepsilon)\right] - (p + \varepsilon) \leq z_j(1 + p), \text{ for all } \varepsilon > 0. \tag{13}$$

Since z_k is continuous and (13) holds for all $\varepsilon > 0$, it follows that

$$\left[1 + p - \sum_{k \neq j} z_k(1 + p)\right] - p \leq z_k(1 + p). \tag{14}$$

Combining (12) and (14) gives

$$1 = \sum_{i \in \bar{U}} z_i(1 + p).$$

Thus by (9),

$$u_i(z_i(1 + p)) = u_i(z_i(1 + p) + p)\delta, \text{ for all } i = 1, \ldots, n.$$

By Lemma 5, this yields $z_i(1 + p) = z_i^*$ for all i, and $p = p^*$.

If: Let all U bargainers bid $p = p^*$. By construction, $z_i(1 + p^*) = z_i^*$ for all $i \in \bar{U}$. We show this does constitute an equilibrium. Since $n > 1$ and

$$1 = \sum_{i \in \bar{U}} z_i(1 + p^*), \tag{15}$$

it follows that

$$\left[1 + p^* - \sum_{j \neq i} z_j(1 + p^*)\right] - p^* = z_i(1 + p^*).$$

Thus *decreasing* one's bid does not have payoff consequences. *Increasing* one's bid by $\varepsilon > 0$ is strictly profitable if

$$\left[1 + p^* + \varepsilon - \sum_{j \neq i} z_j(1 + p^* + \varepsilon)\right] - (p^* + \varepsilon) > z_i(1 + p^*).$$

That is, by (15),

$$1 - \sum_{j \neq i} z_j(1 + p^* + \varepsilon) > 1 - \sum_{j \neq i} z_j(1 + p^*),$$

which is in conflict with Lemma 3. Thus all players bidding p^* does constitute an equilibrium. ■

Proof of Proposition 4 Since $\bar{u}(\bar{v}(z_u)) = \bar{u}(z_u)\delta$ and (8) imply $u(\lambda_u^{-1}\bar{v}(z_u)) = u(\lambda_u^{-1}z_u)\delta$ and the definition of v implies $u(\lambda_u^{-1}z_u)\delta = u(v(\lambda_u^{-1}z_u))$ we have $\bar{v}(z_u) = \lambda_u v(\lambda_u^{-1}z_u)$. Thus (10) and (11) can be written

$$u(\lambda_u^{-1}z_u^*)\delta = u\big(\lambda_u^{-1}(z_u^* + p^*)\big), \text{ for all } u \in U,$$
$$\textstyle\sum_{u\in U} z_u^* = 1.$$

By Lemma 4, this characterizes the equilibrium. Letting $y_u^* = \lambda_u^{-1}z_u^*$ for all u, and $d^* = \lambda^{-1}p^*$, this transforms into

$$u(y_u^*) = u\big(y_u^* + d^*\big)\delta,$$
$$\textstyle\sum_{u\in U} y_u^*\lambda_u = 1.$$

Constructing x^* as in (5) now gives the result. ■

Note that, given u_i, there is a function v_i that specifies the present consumption value of x_i in date 1 such that

$$u_i(v_i(x_i)) = u_i(x_i)\delta, \qquad \text{for all } x_i \in [0,1]. \tag{16}$$

By the concavity of u_i, $v_i^{-1}(x_i) - x_i$ is a continuous and monotonically increasing function of x_i.

Proof of Lemma 1 Recall that $c_i > 0$ for all i and $Y \geq 0$. Thus, the function $e_i(\cdot)$ defined by

$$e_i(x_i) := \frac{v_i^{-1}(c_i x_i)}{c_i} - x_i, \text{ for any } x_i \geq 0, \tag{17}$$

is continuous and monotonically increasing.

Define $\bar{e}_i \in (0, \infty]$ by

$$\sup_{x_i \geq 0} e_i(x_i) := \bar{e}_i.$$

Since $e_i(\cdot)$ is continuous and monotonically increasing, also its inverse

$$x_i(e) := e_i^{-1}(e), \text{ for all } e \in [0, \bar{e}_i],$$

is continuous and monotonically increasing in its domain $[0, \bar{e}_i]$. Condition (17) can now be stated in the form

$$x_i(e) = \frac{v_i(c_i(x_i(e) + e))}{c_i}, \text{ for all } e \in [0, \bar{e}_i]. \tag{18}$$

Moreover, since $0 = x_i(0)$ and $\infty = x_i(\bar{e}_i)$, there is, by the Intermediate Value Theorem, a unique $d > 0$ such that

$$\sum_{i=1}^{n} x_i(d) = Y.$$

■

Proof of Proposition 2 Only if: In a stationary SPE the game ends in finite time. Assume that it never ends. Then each player receives zero. This means that in all subgames each player must get zero. Otherwise there would be a subgame where some offer $y = (y_1, \ldots, y_n)$ is accepted. Because of stationarity this offer is accepted in every subgame. In particular, player 1 can deviate in the first period and offer $y = (y_1, \ldots, y_n)$. This is a profitable deviation and constitutes a contradiction with the assumption that there is a stationary SPE where the game never ends.

Assume next that there is a stationary SPE where an offer $x(i)$ by some player $i \in \{1, 2, \ldots, n\}$, is not accepted immediately. Denote by $z(i)$ the equilibrium outcome in a subgame that starts with an offer $x(i)$ of player i. But now player i could offer $z(i)$ instead of $x(i)$; everyone else would accept the offer as in the stationary equilibrium acceptance depends only on the offer.

Thus, in any equilibrium, $i(t)$'s offer $x(i(t)) = (x_j(i(t)))_{j \in N}$ is accepted at stage $t \in \{0, 1, 2, ..\}$. In stationary equilibrium the time index t can be relaxed from $x(i(t))$. An offer x by i is accepted by all $j \neq i$ if

$$x_j(i) \geq v_j(x_j(j)), \text{ for all } j \neq i. \tag{19}$$

Player i's equilibrium offer $x(i)$ maximizes his payoff with respect to constraint (19) and the resource constraint. By A3, all constraints in (19) and the resource constraint must bind. That is,

$$v_j(x_j(i)) = v_j(x_j(j)), \text{ for all } j \neq i, \tag{20}$$

and

$$\sum_{i=1}^{n} x_i(j) = X, \text{ for all } j. \tag{21}$$

Since player i's acceptance decision is not dependent on the name of the proposer, there is $x_i > 0$ such that $x_i(j) = x_i$ for all $j \neq i$. By (20), $x_j(i) < x_j(j)$ for all j. Hence there is $d > 0$ such that

$$\sum_{i=1}^{n} x_i = X - d. \tag{22}$$

By (20) and (22), x and d do meet (1) and (2). Since 1 is the first proposer, the resulting outcome is $x(1) = (x_1 + d, x_2, \ldots, x_n)$.

If: Let x and d meet (1) and (2). Construct the following stationary strategy: Player i always offers x_{-i} and does not accept less than x_i. Player i's offer y is accepted by all $j \neq i$ only if

$$y_j \geq v_j(X - \sum_{k \neq j} x_k) = v_j(x_j + e_1(x_1)), \text{ for all } j \neq i. \tag{23}$$

Since v_j is increasing, and since

$$x_j = v_j(x_j + d), \text{ for all } j \neq i,$$

i's payoff maximizing offer to each j is x_j. ■

Proof of Proposition 2:

Lemma 5 *For any n, there are unique $y(n) \in \mathbb{R}^{\ltimes}$ and $d(n) > 0$ such that*

$$y_i(n) = v_i(y_i(n) + d(n)), \quad \textit{for all } i = 1, \ldots, n, \tag{24}$$

$$\sum_{u=1}^{n} y_i(n) = n - d(n). \tag{25}$$

Proof By Lemma 1. ■

By Lemma 2, the set of allocations the planner can implement under n agents is

$$\left\{ x : \frac{1}{n} \sum_{i=1}^{n} x_i \leq 1, \quad \text{and } x_i \geq y_i(n), \text{ for all } i = 1, \ldots, n \right\}.$$

Lemma 6 *Let $y(n)$ and $d(n)$ be defined as in Lemma 5. Then there is $y^* \in \mathbb{R}^{\mathbb{U}}$ and $d^* > 0$ such that $y_i(n) \to_n y_u^*$, for all $u_i = u$ and $u \in U$, and $d(n) \to_n d^*$, where*

$$y_u^* = v(y_u^* + d^*), \text{ for all } u \in U, \tag{26}$$

$$\sum_{u \in U} \lambda_u y_u^* = 1. \tag{27}$$

Proof By Lemma 5, for any $n = 1, 2, \ldots,$

$$y_i(n) = v_i(y_i(n) + d(n)), \text{ for all } i = 1, \ldots, n, \tag{28}$$

$$\sum_{i=1}^{n} y_i(n) = n - d(n). \tag{29}$$

Dividing both sides of (29) by n,

$$\frac{1}{n} \sum_{i=1}^{n} y_i(n) = 1 - \frac{d(n)}{n}. \tag{30}$$

Define a function $i : U \to \{1, 2, \ldots\}$ such that $u_{i(u)} = u$, for all $u \in U$. By stationarity, $y_{i(u)}(n) = y_j(n)$ if $u = u_j$. The left hand side of (30) can now be written

$$\frac{1}{n}\sum_{i=1}^{n} y_i(n) = \frac{1}{n}\sum_{u\in U} y_{i(u)}(n)\sum_{i=1}^{n} 1_{(u_i=u)}.$$

By the law of large numbers,

$$\lim_{n}\frac{1}{n}\sum_{i=1}^{n} 1_{(u_i=u)} = \lambda_u. \tag{31}$$

Take any subsequence $\{n'\}$ under which $\lim_{n'} y_{i(u)}(n')$ for all u and $\lim_{n'} d(n')$ exist (the limit can be either finite or infinite). Then (28) can be written

$$\lim_{n'} \sum_{u\in U}\lambda_u y_{i(u)}(n') = 1 - \lim_{n'\to\infty}\frac{d(n')}{n'} \tag{32}$$

By (28) $\lim_{n'} d(n') = \infty$ if and only if $\lim_{n'} y_{i(u)}(n') = \infty$ for all u. Thus, by (32), it must be the case that $\lim_{n'} y_{i(u)}(n') = y_u^*$ and $d(n) = d^*$, for some $(y^*, d^*) \in \mathbb{R}_{++}^{|\mathbb{U}|} \times \mathbb{R}_{++}$. By (28), (32) becomes

$$\lim_{n'} \sum_{u\in U}\lambda_u y_{i(u)}(n') = \sum_{u\in U}\lambda_u y_u^* = 1. \tag{33}$$

By Lemma 1 and (28), $y_{i(u)}^*$ is the limit of any converging subsequence $\{y_{i(u)}(n'')\}$, and d^* is the limit of any converging subsequence $\{d(n'')\}$. Thus (y^*, d^*) is the unique limit and by (33), continuity, and (28) it meets the conditions imposed by the lemma. ■

Proposition 5 *As $n \to \infty$, allocation x is implementable by the planner if and only if $x = x^*$.*

Proof Again, define a function $i : U \to \{1, 2, \ldots\}$ such that $u_{i(u)} = u$, for all $u \in U$. By stationarity, $y_{i(u)}(n) = y_j(n)$ if $u = u_j$, for all $j = 1, \ldots, n$. The set of implementable allocations can be written

$$\left\{x \in \mathbb{R}_+^n : \frac{1}{n}\sum_{i=1}^{n} x_i \le 1, \quad \text{and } x_i \ge y_i(n), \text{ for all } i = 1, \ldots, n\right\}$$

$$= \left\{x \in \mathbb{R}_+^n : \frac{1}{n}\sum_{u\in U} x_{i(u)}\sum_{i=1}^{n} 1_{(u_i=u)} \le 1, \quad \text{and } x_j = x_{i(u)} \ge y_{i(u)}(n) \text{ if } u = u_j, \text{ for all } u \in U\right\}$$

Taking the limit,

$$\lim_{n}\left\{x \in \mathbb{R}_+^n : \frac{1}{n}\sum_{u\in U} x_{i(u)}\sum_{i=1}^{n} 1_{(u_i=u)} \le 1, \quad \text{and } x_j = x_{i(u)} \ge y_{i(u)}(n) \text{ if } u = u_j, \text{ for all } u \in U\right\}$$

$$= \left\{x \in \mathbb{R}_+^\infty : \sum_{u\in U} x_{i(u)}\lambda_u \le 1, \quad \text{and } x_j = x_{i(u)} \ge y_{i(u)}^* \text{ if } u = u_j, \text{ for all } u \in U\right\}.$$

By (27), this reduces to

$$\left\{x \in \mathbb{R}_+^\infty : x_j = y^*_{i(u)} \text{ if } u = u_j, \text{ for all } u \in U\right\},$$

which is a singleton $\{x^*\}$, as required by the proposition. ■

References

Baron, D. P. (1989). A noncooperative theory of legislative coalitions. *American Journal of Political Science*, *33*, 1048–1084.

Baron, D. P., & Ferejohn, J. A. (1989). Bargaining in legislatures. *American Political Science Review*, *83*, 1181–1206.

Baron, D. P. (1993). Government formation and endogenous parties. *American Political Science Review*, *87*, 34–47.

Besley, T., & Coate, S. (1997). An economic model of representative democracy. *Quarterly Journal of Economics*, *108*, 85–114.

Binmore, K., Rubinstein, A., & Wolinsky, A. (1986). The Nash bargaining solution in economic modelling. *Rand Journal of Economics*, *17*, 176–188.

Cai, H. (2002). Bargaining on behalf of a constituency. *Journal of Economic Theory*, *92*, 234–273.

Cai, H., & Cont, W. (2004). Agency problems and commitment in delegated Bargaining. *Journal of Economics and Management Strategy*, *13*, 703–729.

Duggan, J., & Cho, S. -J. (2003). Uniqueness of stationary equilibria in a one-dimensional model of bargaining. *Journal of Economic Theory*, *113*, 118–130.

Eraslan, H. (2002). Uniqueness of stationary equilibrium payoffs in the Baron-Ferejohn model. *Journal of Economic Theory*, *103*, 11–30.

Fishburn, P., & Rubinstein, A. (1982). Time preference. *International Economic Review*, *23*, 677–695.

Herings J. -J., & Predtetchinski, A. (2008). One-dimensional bargaining with Markov recognition probabilities, METEOR Research Memoranda 044.

Herrero, M. (1985). *Strategic theory of market institutions*. Unpublished Ph.D Dissertation, LSE.

Kalandrakis, T. (2006). Proposal rights and political power. *American Journal of Political Science*, *50*, 441–448.

Krishna, V., & Serrano, R. (1996). Multilateral bargaining. *Review of Economic Studies*, *63*, 61–80.

Kultti, K., & Vartiainen, H. (2007a). Von Neumann-Morgenstern stable set bridges time-preferences to the Nash solution. *B.E. Journal of Theoretical Economics*, *7*(1), Article 41.

Kultti, K., & Vartiainen, H. (2007b). Bargaining with many players: A limit result. *Economcis Letters*, *101*, 249–252.

Luce, R., & Raiffa, H. (1957). *Games and decisions: Introduction and critical survey*. New York: Wiley.

Morelli, M. (2004). Party formation and policy outcomes under different electoral systems. *Review of Economic Studies*, *71*, 829–853.

Nurmi, H. (1999). *Voting paradoxes and how to deal with them*. Berlin: Springer.

Thomson, W., & Lensberg, T. (1989). *Axiomatic theory of bargaining with a variable number of agents*. Cambridge: Cambridge University Press.

Coalitions and Catastrophic Climate Change

Norman Schofield

1 Introduction

In this essay I shall consider equilibrium models of the political economy, and attempt a discussion of why equilibrium may collapse into chaos. One motivation is to attempt to come to grips with the fact of climate change in order to evaluate whether theories of social choice give us reason to believe that we may be able to avoid future catastrophe.

Since the early work of Hardin (1968) the "tragedy of the commons" has been recognised as a global prisoner' dilemma. In such a dilemma no agent has a motivation to provide for the collective good. In the context of the possibility of climate change, the outcome is the continued emission of greenhouses gases like carbon dioxide into the atmosphere and the acidification of the oceans. There has developed an extensive literature on the *n*-person prisoners' dilemma in an attempt to solve the dilemma by considering mechanisms that would induce cooperation.[1]

The problem of cooperation has also provided a rich source of models of evolution, building on the early work by Trivers (1971, 1985) and Hamilton (1964); Hamilton (1970). Nowak (2011) provides an overview of the recent developments.

Current work on climate change has focussed on how we should treat the future. For example Stern (2007, 2009), Collier (2010) and Chichilnisky (2009a, b) argue essentially for equal treatment of the present and the future. Dasgupta (2005)

This chapter is based on work supported by NSF grant 0715929.

[1] See for example Hardin (1971, 1982), Taylor (1976, 1982), Axelrod and Hamilton (1981), Axelrod (1981); Axelrod (1984), Kreps et al. (1982), Margolis (1982).

N. Schofield (✉)
Center in Political Economy, Washington University in Saint Louis, 1 Brookings Drive, Saint Louis, MO 63130, USA
e-mail: schofield.norman@gmail.com

M. J. Holler and H. Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*,
DOI: 10.1007/978-3-642-35929-3_39, © Springer-Verlag Berlin Heidelberg 2013

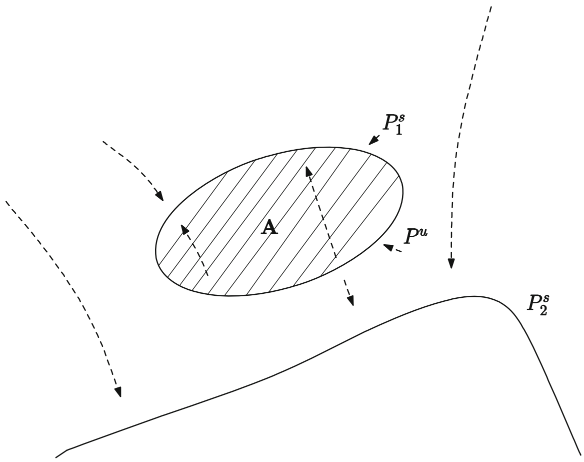


Fig. 1 Stable and unstable components of the global Pareto Set

points out that how we treat the future depends on our current estimates of economic growth in the near future.

The fundamental problem of climate change is that the underlying dynamic system is extremely complex, and displays many positive feedback mechanisms.[2] The difficulty can perhaps be illustrated by Fig. 1. It is usual in economic analysis to focus on Pareto optimality. Typically in economic theory, it is assumed that preferences and production possibilities are generated by convex sets. However, climate change could create non-convexities. In such a case the Pareto set will exhibit stable and unstable components. Figure 1 distinguishes between a domain A, bounded by stable and unstable components P_1^s and P^u, and a second stable component P_2^s. If our actions lead us to an outcome within A, whether or not it is Paretian, then it is possible that the dynamic system generated by climate could lead to a catastrophic destruction of A itself. More to the point, our society would be trapped inside A as the stable and unstable components merged together.

Our society has recently passed through a period of economic disorder, where “black swan” events, low probability occurrences with high costs, have occurred with some regularity. Recent discussion of climate change has also emphasized so called “fat-tailed climate events” again defined by high uncertainty and cost.[3] The catastrophic change implied by Fig. 1 is just such a black swan event. The point to note about Fig. 1 is everything would appear normal until the evaporation of A.

Cooperation between nations to deal with climate change could in principle be attained by the action of a hegemonic leader such as the United States as suggested by Kindleberger (1973) and Keohane and Nye (1977). In Sect. 2 we give a brief exposition of the prisoners’ dilemma and illustrate how hegemonic behavior could

[2] See the discussion in Schofield (2011).

[3] Weitzman (2009) and Chichilnisky (2010). See also Chichilinisky ((1996b, 2000, 2011a, b) and Eisenberger (2010) on other catastrophic events such as collision with an asteroid.

facilitate international cooperation. However, the analysis suggests that in the present economic climate, such hegemonic leadership is unlikely.

Analysis of games such as the prisoner's dilemma usually focus on the existence of a Nash equilibrium, a vector of strategies with the property that no agent has an incentive to change strategy. Section 3 considers the family of equilibrium models based on the Brouwer (1912) fixed point theorem, or the more general result known as the Ky Fan theorem (Fan 1961) as well as the application by Bergstrom (1975, 1992) to prove existence of a Nash equilibrium and market equilibrium.

Section 4 considers a generalization of the Ky Fan Theorem, and argues that the general equilibrium argument can be interpreted in terms of particular properties of a preference field, H, defined on the tangent space of the joint strategy space. If this field is continuous, in a certain well-defined sense, and "*half open*" then it will exhibit a equilibrium. This half open property is the same as the non empty intersection of a family of dual cones. We mention a Theorem by Chichilnisky (1995) that a necessary and sufficient condition for market equilibrium is that a family of dual cones also has non-empty intersection.

However, preference fields that are defined in terms of coalitions need not satisfy the half open property and thus need not exhibit equilibrium. For coalition systems, it can be shown that unless there is a collegium or oligarchy, or the dimension of the space is restricted in a particular fashion, then there need be no equilibrium. Earlier results by McKelvey (1976), Schofield (1978), McKelvey and Schofield (1987) and Saari (1995, 1997) suggested that voting can be "non-equilibrating" and indeed "chaotic."[4]

Kauffman (1993) commented on "chaos" or the failure of "structural stability" in the following way.

> One implication of the occurrence or non-occurrence of structural stability is that, in structurally stable systems, smooth walks in parameter space must [result in] smooth changes in dynamical behavior. By contrast, chaotic systems, which are not structurally stable, adapt on uncorrelated landscapes. Very small changes in the parameters pass through many interlaced bifurcation surfaces and so change the behavior of the system dramatically.

Chaos is generally understood as sensitive dependence on initial conditions whereas *structural stability* means that the qualitative nature of the dynamical system does not change as a result of a small perturbation.[5] I shall use the term *chaos* to mean that the trajectory taken by the dynamical process can wander anywhere.[6]

---

[4] See Schofield (1977); Schofield (1980a); Schofield (1980b). In a sense these voting theorems can be regarded as derivative of Arrow's Impossibility Theorem (Arrow 1951). See also Arrow (1986).

[5] The theory of chaos or complexity is rooted in Smale's fundamental theorem (Smale 1966) that structural stability of dynamical systems is not "generic" or typical whenever the state space has more than two dimensions.

[6] In their early analysis of chaos, Li and Yorke (1975) showed that in the domain of a chaotic transformation f it was possible for almost any pair of positions (x, y) to transition from x to $y = f^r(x)$, where f^r means the r times reiteration of f.

An earlier prophet of uncertainty was, of course, Keynes (1936) whose ideas on "speculative euphoria and crashes" would seem to be based on understanding the economy in terms of the qualitative aspects of its coalition dynamics.[7] An extensive literature has tried to draw inferences from the nature of the recent economic events. A plausible account of market disequilibrium is given by Akerlof and Shiller (2009) who argue that

> the business cycle is tied to feedback loops involving speculative price movements and other economic activity — and to the talk that these movements incite. A downward movement in stock prices, for example, generates chatter and media response, and reminds people of longstanding pessimistic stories and theories. These stories, newly prominent in their minds, incline them toward gloomy intuitive assessments. As a result, the downward spiral can continue: declining prices cause the stories to spread, causing still more price declines and further reinforcement of the stories.

It would seem reasonable that the rise and fall of the market is due precisely to the coalitional nature of decision-making, as large sets of agents follow each other in expecting first good things and then bad. A recent example can be seen in the fall in the market after the earthquake in Japan, and then recovery as an increasing set of investors gradually came to believe that the disaster was not quite as bad as initially feared.

Since investment decisions are based on these uncertain evaluations, and these are the driving force of an advanced economy, the flow of the market can exhibit singularities, of the kind that recently nearly brought on a great depression. These singularities associated with the bursting of markets bubbles are time-dependent, and can be induced by endogenous belief-cascades, rather than by any change in economic or political fundamentals (Corcos et al. 2002).

Similar uncertainty holds over political events. The fall of the Berlin Wall in 1989 was not at all foreseen. Political scientists wrote about it in terms of "belief cascades"[8] as the coalition of protesting citizens grew apace. As the very recent democratic revolutions in the Middle East and North Africa suggest, these coalitional movements are extremely uncertain.[9] In particular, whether the autocrat remains in power or is forced into exile is as uncertain as anything Keynes discussed. Even when democracy is brought about, it is still uncertain whether it will persist.[10]

Section 5 introduces the Condorcet (1994, [1795]) Jury Theorem. This theorem suggests that majority rule can provide a way for a society to attain the truth when the individuals have common goals. Schofield (2002, 2006) has argued that Madison was aware of this theorem while writing Federalist X (Madison 1999, [1787]) so it can be taken as perhaps the ultimate justification for democracy.

---

[7] See Minsky (1975); Minsky (1986).

[8] Karklins and Petersen (1993), Lohmann (1994). See also Bikhchandani et al. (1992).

[9] The response by the citizens of these countries to the demise of Osama bin Laden on May 2, 2011, is in large degree also unpredictable.

[10] See for example Carothers (2002) and Collier (2009).

However, models of belief aggregation that are derived from the Jury Theorem can lead to belief cascades that bifurcate the population. In addition, if the aggregation process takes place on a network, then centrally located agents, who have false beliefs, can dominate the process.[11]

In Sect. 6 we introduce the idea of a belief equilibrium, and then go on to consider the notion of "punctuated equilibrium" in general evolutionary models. Again however, the existence of an equilibrium depends on a fixed point argument, and thus on a half open property of the "cones" by which the developmental path is modeled. This half open property is equivalent to the existence of a social direction gradient defined everywhere. In Sect. 7 we discuss René Thom's notion of a structurally stable "chreod" and introduce the heart of a social process. This depends on the idea of localizing the degree of chaos in a dynamical system. Finally we link Popper's assertion that prediction is impossible in the social sciences (Popper 1959) with the failure of acyclicity in general social systems. Section 8 concludes with some remarks on unpredictability in political decision-making.

2 The Prisoners' Dilemma, Cooperation and Morality

> For before constitution of Sovereign Power ... all men had right to all things; which necessarily causeth Warre. Hobbes (2009, [1651]).

Kindleberger (1973) gave the first interpretation of the international economic system of states as a "Hobbesian" prisoners' dilemma, which could be solved by a leader, or "hegemon."

> A symmetric system with rules for counterbalancing, such as the gold standard is supposed to provide, may give way to a system with each participant seeking to maximize its short-term gain. ...But a world of a few actors (countries) is not like [the competitive system envisaged by Adam Smith]. ... In advancing its own economic good by a tariff, currency depreciation, or foreign exchange control, a country may worsen the welfare of its partners by more than its gain. Beggar-thy-neighbor tactics may lead to retaliation so that each country ends up in a worse position from having pursued its own gain ...

> This is a typical non-zero sum game, in which any player undertaking to adopt a long range solution by itself will find other countries taking advantage of it ...

In the 1970s, Keohane and Nye (1977) and Keohane (1984) rejected "realist" theory in international politics, and made use of the idea of a hegemonic power in a context of "complex interdependence" of the kind envisaged by Kindleberger. Although they did not refer to the formalism of the prisoners' dilemma, it would appear that this notion does capture elements of complex interdependence. To

[11] Golub and Jackson (2010).

some extent, their concept of a hegemon is taken from realist theory rather than deriving from the game-theoretic formalism.

However, it is very easy to adapt the notion of a symmetric prisoners' dilemma, so as to clarify the concept of a hegemon. A non-symmetric n-agent prisoners' dilemma (*nPD*) can be constructed as follows. We define the strategy of the ith country to be $d_i = 0$ when i defects, and $d_i = 1$ when i cooperates. We can also denote mixed strategies by letting $d_i \in [0,1]$. Each country was a weight (proportional to its GDP), a_i say. The *total collective good* of the system, N, of states is defined to be:

$$B(N) = \sum_{j=1}^{n} a_j d_j \tag{1}$$

The payoff u_i to state i, when it adopts strategy d_i is

$$u_i(d_i) = \frac{r}{n} B(N) - d_i. \tag{2}$$

To construct a prisoners' dilemma, we assume $1 < r < n$. From the above, the term involving d_i in $u_i(d_i)$ is

$$\frac{r}{n}(a_i d_i) - d_i. \tag{3}$$

Clearly if

$$\frac{r}{n}(a_i) < 1, \tag{4}$$

then u_i is maximized at $d_i = 0$, and takes the value $\frac{r}{n} B(N) = \frac{r}{n} \sum_{j \neq i} a_j d_j$. If

$$\frac{r}{n}(a_i) > 1, \tag{5}$$

then u_i is maximized at $d_i = 1$.

In the *symmetric* game, $a_i = 1$ for all i, so the "rational" strategy for each country is to defect, by choosing $d_i = 0$. In this case $u_i = 0$. On the other hand if everyone chooses the *irrational* strategy $d_i = 1$, then $B(N) = r > 1$, and so $u_i(d_i = 1) > 0$.

On the other hand, if

$$a_i > \frac{n}{r}, \tag{6}$$

then $a_i > 1$, and this country, i, rationally must *cooperate*, irrespective of the strategies of other countries. To keep things simple, suppose $a_j = 1$ for all j other than this hegemon, i. In this very trivial formulation, some things are obvious. If more states join the game (so n increases, while r *remains constant*), it becomes more "difficult" for a_i to be large enough for cooperation. The coefficient, r, is the "rate of return on cooperation." As r falls it becomes more difficult again for i to remain the cooperative hegemon. In this formulation the term hegemon is something of a misnomer, since i is simply a rational cooperator. However, if coalitions are possible and a hegemonic power, called i, leads a coalition M of states, dictating policy to these states, then the optimality condition for the joint cooperation of the states in the coalition M is

$$\sum_{i \in M} a_i > \frac{n}{r}. \tag{7}$$

The collective benefits of the coalition M can then be redistributed by the hegemon in some way, to keep the coalition intact. The essence of the theory of hegemony in international relations is that if there is a degree of inequality in the strengths of nation states then a hegemonic power may maintain cooperation in the context of an n-country prisoners' dilemma. Clearly, the British Empire in the 1800s is the role model for such a hegemon (Ferguson 2002). See also Schofield (1975, 1985).

Hegemon theory suggests that international cooperation was maintained after World War II because of a dominant cooperative coalition. At the core of this cooperative coalition was the United States; through its size it was able to generate collective goods for this community, first of all through the Marshall Plan and then in the context first of the post-world war II system of trade and economic cooperation, based on the Bretton Woods agreement and the Atlantic Alliance, or NATO. Over time, the United States has found it costly to be the dominant core of the coalition In particular, as the relative size of the U.S. economy has declined, so that $\sum_{i \in M} a_i$ has fallen, then cooperation will become very difficult, especially if r also falls. Indeed, the global recession of 2008–10 suggests that problems of debt could induce "begger thy neighbor strategies", just like the 1930s.

The future utility benefits of adopting policies to ameliorate these possible changes depend on the discount rates that we assign to the future. Dasgupta (2005) gives a clear exposition of how we might assign these discount rates. Obviously enough, different countries will in all likelihood adopt very different evaluations of the future. Developing countries like the BRICs (Brazil, Russia, India and China) will choose growth and development now rather than choosing consumption in the future.

There have been many attempts to "solve" the prisoners' dilemma in a general fashion. For example Binmore (2005, 2009) suggests that in the iterated nPD there are many equilibria with those that are *fair* standing out in some fashion. However, the criterion of "fairness" would seem to have little weight with regard to climate change. It is precisely the poor countries that will suffer from climate change, while the rapidly growing BRICS believe that they have a right to choose their own paths of development.

An extensive literature over the last few years has developed Smith's ideas as expressed in the *Theory of Moral Sentiments* (1984 [1759]) to argue that human beings have an inate propensity to cooperate. This propensity may well have been the result of co-evolution of language and culture (Boyd and Richerson 2005; Gintis 2000).

Since language evolves very quickly (McWhorter 2001; Deutscher 2006), we might also expect moral values to change fairly rapidly, at least in the period during which language itself was evolving. In fact there is empirical evidence that cooperative behavior as well as notions of fairness vary significantly across different societies.[12] While there may be fundamental aspects of morality and "altruism," in particular, held in common across many societies, there is variation in how these are articulated. Gazzaniga (2008) suggests that moral values can be described in terms of various *modules*: reciprocity, suffering (or empathy), hierarchy, in-group and outgroup coalition, and purity/disgust. These modules can be combined in different ways with different emphases. An important aspect of cooperation is emphasized by Burkhart et al. (2009) and Hrdy (2011), namely cooperation between man and woman to share the burden of child rearing.

It is generally considered that hunter-gatherer societies adopted egalitarian or "fair share" norms. The development of agriculture and then cities led to new norms of hierarchy and obedience, coupled with the predominance of military and religious elites (Schofield 2010).

North (1990), North et al. (2009) and Acemoglu and Robinson (2006) focus on the transition from such oligarchic societies to open access societies whose institutions or "rules of the game", protect private property, and maintain the rule of law and political accountability, thus facilitating both cooperation and economic development. Acemoglu et al. (2009) argue, in their historical analyses about why "good" institutions form, that the evidence is in favor of "critical junctures."[13] For example, the "Glorious Revolution" in Britain in 1688 (North and Weingast 1989), which prepared the way in a sense for the agricultural and industrial revolutions to follow (Mokyr 2005, 2010; Mokyr and Nye 2007) was the result of a sequence of historical contingencies that reduced the power of the elite to resist change. Recent work by Morris (2010), Fukuyama (2011), Ferguson (2011) and Acemoglu and Robinson (2011) has suggested that these fortuitous circumstances never occurred in China and the Middle East, and as a result these domains fell behind the West. Although many states have become democratic in the last few decades, oligarchic power is still entrenched in many parts of the world.[14]

At the international level, the institutions that do exist and that are designed to maintain cooperation, are relatively young. Whether they succeed in facilitating cooperation in such a difficult area as climate change is a matter of speculation. As we have suggested, international cooperation after World War II was only possible because of the overwhelming power of the United States. In a world with

[12] See Henrich et al. (2004, 2005), which reports on experiments in fifteen "small-scale societies," using the game theoretic tools of the "prisoners' dilemma," the "ultimatum game," etc.

[13] See also Acemoglu and Robinson (2008).

[14] The popular protests in N.Africa and the Middle East in 2011 were in opposition to oligarchic and autocratic power.

oligarchies in power in Russia, China, and in many countries in Africa, together with political disorder in almost all the oil producing counties in the Middle East, cooperation would appear unlikely.

To extend the discussion, we now consider more general theories of social choice.

3 Existence of a Choice

The above discussion has considered a very simple version of the prisoner's dilemma. The more general models of cooperation typically use variants of evolutionary game theory, and in essence depend on proof of existence of Nash equilibrium, using some version of the Brouwer's fixed point theorem (Brouwer 1912).

Brouwer's theorem asserts that any continuous function $f : B \to B$ from the finite dimensional ball, B (or indeed any compact convex set in $\mathbb{R}^w$) into itself, has the *fixed point property*. That is, there exists some $x \in B$ such that $f(x) = x$.

We will now consider the use of variants of the theorem, to prove existence of an equilibrium of a general choice mechanism. We shall argue that the condition for existence of an equilibrium will be violated if there are cycles in the underlying mechanism.

Let W be the set of alternatives and let X be the set of all subsets of W. A *preference correspondence*, P, on W assigns to each point $x \in W$, its *preferred set* $P(x)$. Write $P : W \to X$ or $P : W \twoheadrightarrow W$ to denote that the image of x under P is a set (possibly empty) in W. For any subset V of W, the restriction of P to V gives a correspondence $P_V : V \twoheadrightarrow V$. Define $P_V^{-1} : V \twoheadrightarrow V$ such that for each $x \in V$,

$$P_V^{-1}(x) = \{y : x \in P(y)\} \cap V.$$

$P_V^{-1}(x) = \{y : x \in P(y)\} \cap V$. The sets $P_V(x), P_V^{-1}(x)$ are sometimes called the *upper* and *lower* preference sets of P on V. When there is no ambiguity we delete the suffix V. The *choice* of P from W is the set

$$C(W, P) = \{x \in W : P(x) = \Phi\}.$$

Here Φ is the empty set. The choice of P from a subset, V, of W is the set

$$C(V, P) = \{x \in V : P_V(x) = \Phi\}.$$

Call C_P a *choice function* on W if $C_P(V) = C(V, P) \neq \Phi$ for every subset V of W. We now seek general conditions on W and P which are sufficient for C_P to be a choice function on W. Continuity properties of the preference correspondence are important and so we require the set of alternatives to be a topological space.

Definition 1 Let W, Y be two topological spaces. A correspondence $P : W \twoheadrightarrow Y$ is

(i) *Lower demi-continuous* (*ldc*) iff, for all $x \in Y$, the set

$$P^{-1}(x) = \{y \in W : x \in P(y)\}$$

is open (or empty) in W.

(ii) *Acyclic* if it is impossible to find a cycle $x_t \in P(x_{t-1}), x_{t-1} \in P(x_{t-2}), .., x_1 \in P(x_t)$.

(iii) *Lower hemi-continuous* (*lhc*) iff, for all $x \in W$, and any open set $U \subset Y$ such that $P(x) \cap U \neq \Phi$ there exists an open neighborhood V of x in W, such that $P(x') \cap U \neq \Phi$ for all $x' \in V$. Note that if P is ldc then it is lhc.

We shall use lower demi-continuity of a preference correspondence to prove existence of a choice.

We shall now show that if W is compact, and P is an acyclic and ldc preference correspondence $P : W \twoheadrightarrow W$, then $C(W, P) \neq \Phi$. First of all, say a preference correspondence $P : W \twoheadrightarrow W$ satisfies the *finite maximality property* (FMP) on W iff for every finite set V in W, there exists $x \in V$ such that $P(x) \cap V = \Phi$.

Lemma 1(Walker 1977) *If W is a compact, topological space and P is an ldc preference correspondence that satisfies FMP on W, then $C(W, P) \neq \Phi$.*

This follows readily, using compactness to find a finite subcover, and then using FMP.

Corollary 1 *If W is a compact topological space and P is an acyclic, ldc preference correspondence on W, then $C(W, P) \neq \Phi$.*

As Walker (1977) noted, when W is compact and P is ldc, then P is acyclic iff P satisfies FMP on W, and so either property can be used to show existence of a choice. A second method of proof is to show that C_P is a choice function is to substitute a convexity property for P rather than acyclicity.

Definition 2

(i) If W is a subset of a vector space, then the *convex hull* of W is the set, $Con[W]$, defined by taking all convex combinations of points in W.

(ii) W is *convex* iff $W = Con[W]$. (The empty set is also convex.)

(iii) W is *admissible* iff W is a compact, convex subset of a topological vector space.

(iv) A preference correspondence $P : W \twoheadrightarrow W$ is *semi-convex* iff, for all $x \in W$, it is the case that $x \notin Con(P(x))$.

Fan (1961) has shown that if W is admissible and P is ldc and semi-convex, then $C(W, P)$ is non-empty.

Choice Theorem (Fan 1961; Bergstrom 1975)

If W is an admissible subset of a Hausdorff topological vector space, and $P : W \twoheadrightarrow W$ a preference correspondence on W which is ldc and semi-convex then $C(W, P) \neq \Phi$.

The proof uses the KKM lemma due to Knaster et al. (1929).

The original form of the Theorem by Fan made the assumption that $P: W \twoheadrightarrow W$ was *irreflexive* (in the sense that $x \notin P(x)$ for all $x \in W$) and *convex*. Together these two assumptions imply that P is semi-convex. Bergstrom (1975) extended Fan's original result to give the version presented above.[15]

Note that the Fan Theorem is valid without restriction on the dimension of W. Indeed, Aliprantis and Brown (1983) have used this theorem in an economic context with an infinite number of commodities to show existence of a price equilibrium. Bergstrom (1992) also showed that when W is finite dimensional then the Fan Theorem is valid when the continuity property on P is weakened to lhc and used this theorem to show existence of a Nash equilibrium of a game $G = \{(P_1, W_1), .P_i, W_i), ..(P_n, W_n) : i \in N\}$. Here the ith strategy space is a finite dimensional space, W_i and each individual has a preference P_i on the joint strategy space $P_i : W^N = W_1 \ \times W_2 \ldots \times W_n \ \twoheadrightarrow W_i$. The Fan Theorem can be used, in principle to show existence of an equilibrium in complex economies with externalities. Define the Nash improvement correspondence by $P_i^* : W^N \twoheadrightarrow W^N$ by $y \in P_i^*(x)$ whenever $y = (x_1, ..x_{i-1}, x_i^*, \ldots, x_n)$, $x = (x_1, .., x_{i-1}, x_i, .., x_n)$, and $x_i^* \in P_i(x)$ The joint Nash improvement correspondence is $P_N^* = \cup P_i^* : W^N \twoheadrightarrow W^N$. The Nash equilibrium of a game G is a vector $\mathbf{z} \in \mathbf{W^N}$ such that $P_N^*(\mathbf{z}) = \Phi$. Then the Nash equilibrium will exist when P_N^* is ldc and semi-convex and W^N is admissible.

4 Dynamical Choice Functions

We now consider a *generalized preference field* $H : W \twoheadrightarrow TW$, on a manifold W. TW is the tangent bundle above W, given by $TW = \cup\{T_xW : x \in W\}$, where T_xW is the tangent space above x. If V is a neighborhood of x, then $T_VW = \cup\{T_xW : x \in V\}$ which is locally like the product space $\mathbb{R}^w \times V$. Here W is locally like $\mathbb{R}^w$.

At any $x \in W$, $H(x)$ is a *cone* in the tangent space T_xW above x. That is, if a vector $v \in H(x)$, then $\lambda v \in H(x)$ for any $\lambda > 0$. If there is a smooth curve, $c : [-1, 1] \to W$, such that the differential $\frac{dc(t)}{dt} \in H(x)$, whenever $c(t) = x$, then c is called an *integral curve* of H. An integral curve of H from $x = c(o)$ to $y = \lim_{t \to 1} c(t)$ is called an H - *preference curve* from x to y. In this case we write $y \in \mathbb{H}(x)$. We say y is reachable from x if there is a piecewise differentiable H—preference curve from x to y, so $y \in \mathbb{H}^r(x)$ for some reiteration r. The preference field H is called *S-continuous* iff the inverse relation $\mathbb{H}^{-1}$ is ldc. That is, if x is reachable from y, then there is a neighborhood V of y such that x is reachable from all of V. The *choice* $C(W, H)$ of H on W is defined by

$$C(W, H) = \{x \in W : H(x) = \Phi\}.$$

[15] See also Shafer and Sonnenschein (1975).

Say $H(x)$ is semi-convex at $x \in W$, if either $H(x) = \Phi$ or $0 \notin Con[H(x)]$ in the tangent space T_xW. In the later case, there will exist a vector $v' \in T_xW$ such that $(v' \cdot v) > 0$ for all $v \in H(x)$. We can say in this case that there is, at x, a *direction gradient* d in the cotangent space T_x^*W of linear maps from T_xW to $\mathbb{R}$ such that $d(v) > 0$ for all $v \in H(x)$. If H is S-continuous and half-open in a neighborhood, V, then there will exist such a continuous direction gradient $d : V \to T^*V$ on the neighborhood V

We define

$$Cycle(W, H) = \{x \in W : H(x) \neq \Phi, 0 \in \mathrm{Con}H(x)\}.$$

An alternative way to characterize this property is as follows.

Definition 3 The *dual* of a preference field $H : W \twoheadrightarrow TW$ is defined by $H^* : W \twoheadrightarrow T^*W : x \twoheadrightarrow \{d \in T_x^*W : d(v) > 0 \text{ for all } v \in H(x) \subset T_xW\}$. For convenience if $H(x) = \Phi$ we let $H^*(x) = T_xW$. Note that if $0 \notin ConH(x)$ iff $H^*(x) \neq \Phi$. We can say in this case that the field is *half open* at x.[16]

In applications, the field $H(x)$ at x will often consist of some family $\{H_j(x)\}$. As an example , let $u : W \twoheadrightarrow \mathbb{R}^n$ be a smooth utility profile and for any coalition $M \subset N$ let

$$H_M(u)(x) = \{v \in T_xW : (du_i(x)(v) > 0, \forall\, i \in M\}.$$

If $\mathbb{D}$ is a family of *decisive* coalitions, $\mathbb{D} = \{M \subset N\}$, then we define

$$H_{\mathbb{D}}(u) = \cup H_M(u) : W \twoheadrightarrow TW$$

Then the field $H_{\mathbb{D}}(u) : W \twoheadrightarrow TW$ has a dual $[H_{\mathbb{D}}(u)]^* : W \twoheadrightarrow T^*W$ given by $[H_{\mathbb{D}}(u)]^*(x) = \cap[H_M(u)(x)]^*$ where the intersection at x is taken over all $M \in \mathbb{D}$ such that $H_M(u)(x) \neq \Phi$. We call $[H_M(u)(x)]^*$ the *co-cone of* $[H_M(u)(x)]^*$. It then follows that at $x \in Cycle(W, H_{\mathbb{D}}(u))$ then $0 \in Con[H_{\mathbb{D}}(u)(x)]$ and so $[H_{\mathbb{D}}(u)(x)]^* = \Phi$. Thus

$$Cycle(W, H_{\mathbb{D}}(u)) = \{x \in W : [H_{\mathbb{D}}(u)]^*(x) = \Phi\}.$$

The condition that $[H_{\mathbb{D}}(u)]^*(x) = \Phi$ is equivalent to the condition that $\cap[H_M(u)(x)]^* = \Phi$ and was called the *null dual condition* (at x). Schofield (1978) has shown that $Cycle(W, H_{\mathbb{D}}(u))$ will be an open set and contains cycles so that a point x is reachable from itself through a sequence of preference curves associated with different coalitions. This result was an application of a more general result.

Dynamical Choice Theorem (Schofield 1978)

For any S-continuous field H on compact, convex W, then

$$Cycle(W, H) \cup C(W, H) \neq \Phi.$$

---

[16] In other words, there exists d such that $d(v) > 0$ for all $v \in H(x) \subset T_xW$, whenever $H(x) \neq \Phi$.

If $x \in Cycle(W,H) \neq \Phi$ then there is a *piecewise differentiable* H *-preference cycle* from x to itself. If there is an open path connected neighborhood $V \subset Cycle(W,H)$ such that $H(x')$ is open for all $x' \in V$ then there is a *piecewise differentiable* H *-preference curve* from x to x'. □

(Here piecewise differentiable means the curve is continuous, and also differentiable except at a finite number of points). The proof follows from the previous choice theorem. The trajectory is built up from a set of vectors $\{v_1, \ldots v_t\}$ each belonging to $H(x)$ with $0 \in Con[\{v_1, \ldots v_t\}]$. If $H(x)$ is of full dimension, as in the case of a voting rule, then just as in the model of chaos by Li and Yorke (1975), trajectories defined in terms of H can wander anywhere within any open path connected component of $Cycle(W,H)$.

This result has been shown more generally in Schofield (1984a) for the case that W is a compact manifold with non-zero Euler characteristic (Brown 1971). For example the theorem is valid if W is an even dimensional sphere. (The theorem is not true on odd dimensional spheres, as the clock face illustrates.)

Existence of Nash Equilibrium

Let $\{W_1, \ldots, W_n\}$ be a family of compact, contractible, smooth, strategy spaces with each $W_i \subset \mathbb{R}^w$. A smooth profile $u{:}W^N = W_1 \times W_2 \ldots \times W_n \twoheadrightarrow \mathbb{R}^n$. Let $H_i : W_i \twoheadrightarrow TW_i$ be the induced i -preference field in the tangent space over W_i. If each H_i is S-continuous and half open in TW_i then there exists a *critical Nash equilibrium,* $\mathbf{z} \in \mathbf{W^N}$ such that $H^N(\mathbf{z}) = (H_1 \times ..H_n)(\mathbf{z}) = \Phi$.

This follows from the choice theorem because the product preference field, H^N, will be half-open and S-continuous. Below we consider existence of *local* Nash equilibrium.[17] With smooth utility functions, a local Nash equilibrium can be found by checking the second order conditions on the Hessians. (See Schofield 2007, for an application of this technique.)

Example 1 To illustrate the Choice Theorem, define the preference relation $P_{\mathbb{D}} : W \twoheadrightarrow W$ generated by a family of *decisive* coalitions, $\mathbb{D} = \{M \subset N\}$, so that $y \in P_{\mathbb{D}}(x)$ whenever all voters in some coalition $M \in \mathbb{D}$ prefer y to x. In particular consider the example due to Kramer (1973), with $N = \{1,2,3\}$ and $\mathbb{D} = \{\{1,2\},\{1,3\},\{2,3\}$. Suppose further that the preferences of the voters are characterized by the direction gradients

$$\{du_i(x) : i = 1,2,3\}$$

as in Fig. 2. In the figure, the utilities are assume to be "Euclidean," derived from distance from a preferred point, but this assumption is not important.

As the figure makes evident, it is possible to find three points $\{a,b,c\}$ in W such that

[17] That is a critical Nash equilibrium which is an attractor of the integral curves.

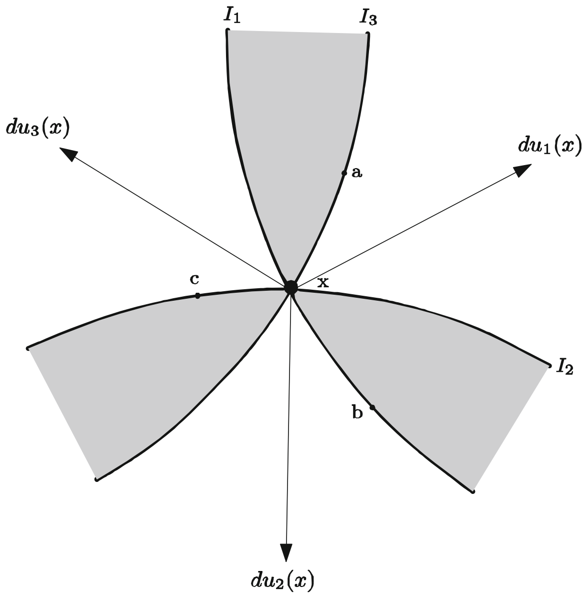


Fig. 2 Cycles in a neighborhood of x

$$\begin{aligned} u_1(a) &> u_1(b) = u_1(x) > u_1(c) \\ u_2(b) &> u_2(c) = u_2(x) > u_2(a) \\ u_3(c) &> u_3(a) = u_3(x) > u_3(b). \end{aligned}$$

That is to say, preferences on $\{a, b, c\}$ give rise to a *Condorcet cycle*. Note also that the set of points $P_{\mathbb{D}}(x)$, preferred to x under the voting rule, are the shaded "win sets" in the figure. Clearly $x \in ConP_{\mathbb{D}}(x)$, so $P_{\mathbb{D}}(x)$ is not semi-convex. Indeed it should be clear that in *any* neighborhood V of x it is possible to find three points $\{a', b', c'\}$ such that there is *local* voting cycle, with $a' \in P_{\mathbb{D}}(b'), b' \in P_{\mathbb{D}}(c'), c' \in P_{\mathbb{D}}(a')$. We can write this as

$$a' \to c' \to b' \to a'.$$

Not only is there a voting cycle, but the Fan theorem fails, and we have no reason to believe that $C(W, P_{\mathbb{D}}) \neq \Phi$.

We can translate this example into one on preference fields by considering the preference field

$$H_{\mathbb{D}}(u) = \cup H_M(u) : W \twoheadrightarrow TW$$

where each $M \in \mathbb{D}$.

Figure 3 shows the three different preference fields $\{H_i : i = 1, 2, 3)$ on W, as well as the intersections H_M, for $M = \{1, 2\}$ etc.

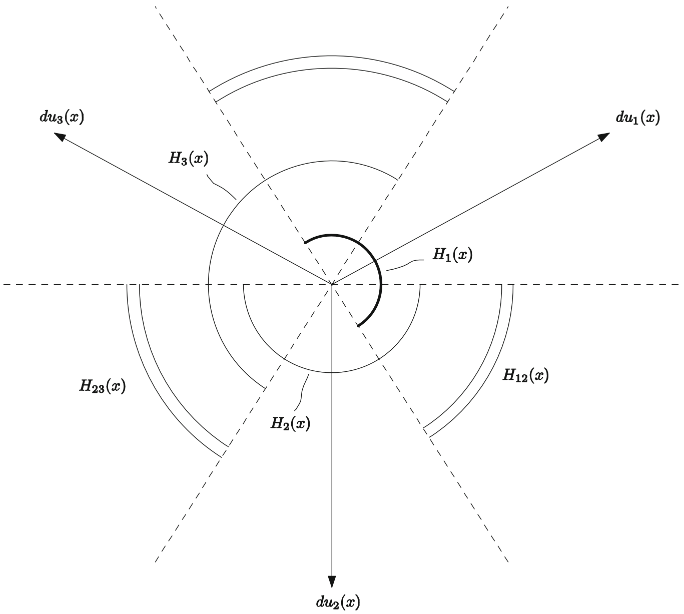


Fig. 3 The failure of half-openness of a preference field

Obviously the joint preference field $H_{\mathbb{D}}(u) = \cup H_M(u) : W \twoheadrightarrow TW$ fails the half open property at x since $0 \in Con[H_{\mathbb{D}}(u)(x)]$. Although $H_{\mathbb{D}}(u)$ is S-continuous, we cannot infer that $C(W, H_{\mathbb{D}}(u)) \neq \Phi$.

Chichilnisky (1992, 1995, 1996a, 1997a) has obtained similar results for markets, where the condition that the dual is non-empty was termed *market arbitrage,* and defined in terms of global market co-cones associated with each player. Such a dual co-cone, $[H_i(u)]^*$ is precisely the set of prices in the cotangent space that lie in the dual of the prefered cone, $[H_i(u)]$, of the agent. By analogy with the above, she identifies this condition on non-emptiness of the intersection of the family of co-cones as one which is necessary and sufficient to guarantee an equilibrium.

Chichilnisky Theorem (Chichilnisky 1997b)
The *limited arbitrage condition* $\cap[H_i(u)]^* \neq \Phi$ is necessary and sufficient for existence of a competitive equilibrium.□

Chichilnisky (1993, 1997c) also defined a topological obstruction to the non-emptiness of this intersection and showed the connection with the existence of a social choice equilibrium.

For a voting rule, $\mathbb{D}$ it is possible to guarantee that $Cycle(W, H_{\mathbb{D}}) = \Phi$ and thus that $C(W, HD) \neq \Phi$. We can do this by restricting the dimension of W.

Definition 4

(i) Let $\mathbb{D}$ be a family of decisive subsets of the finite society N of size n. If the collegium, $K(\mathbb{D}) = \cap\{\mathbb{M} \in \mathbb{D}\}$ is non-empty then $\mathbb{D}$ is called *collegial* and the *Nakamura number* $\kappa(\mathbb{D})$ is defined to be ∞.
(ii) If the collegium $K(\mathbb{D})$ is empty then $\mathbb{D}$ is called *non-collegial*. Define the *Nakamura number* in this case to be $\kappa(\mathbb{D}) = \min\{|\mathbb{D}'| \nLeftarrow \mathbb{D}' \subset \mathbb{D}$ and $K(\mathbb{D}') = \Phi\}$.

Nakamura Theorem
If $u \in U(W)^N$ and $\mathbb{D}$ has Nakamura number $\kappa(\mathbb{D})$ with $dim(W) \leq \kappa(\mathbb{D}) - 2$ then $Cycle(W, H_{\mathbb{D}}(u)) = \Phi$ and $C(W, H_{\mathbb{D}}(u)) \neq \Phi$.

Outline of proof. Consider any subfamily $\mathbb{D}'$ of $\mathbb{D}$ with cardinality $\kappa(\mathbb{D}) - 1$. Then $\cap M \neq \Phi$, so $\cap\{[H_M(u)]^*(x) : M \in \mathbb{D}'\} \neq \Phi$. If $[H_M(u)(x)] \neq \Phi$, we can identify each $[H_M(u)(x)]^*$ with a non-empty convex hull generated by $(du_i(x) : i \in M\}$. These sets can be projected into T_xW where they are convex and compact. Since $dim(W) \leq \kappa(\mathbb{D}) - 2$, then by Helly's Theorem, we see that $\cap\{[H_M(u)]^*(x) : M \in \mathbb{D}\} \neq \Phi$. Thus $Cycle(W, H_{\mathbb{D}}(u)) = \Phi$ and $C(W, H_{\mathbb{D}}(u)) \neq \Phi$. □

See Schofield (1984b), Nakamura (1979) and Strnad (1985).

For social choice defined by voting games, the Nakamura number for majority rule is 3, except when $n = 4$, in which case $\kappa(\mathbb{D}) = 4$, so the Nakamura Theorem can generally only be used to prove a "median voter" theorem in one dimension. However, the result can be combined with the Fan Theorem to prove existence of equilibrium for a political economy with voting rule $\mathbb{D}$, when the dimension of the public good space is no more than $\kappa(\mathbb{D}) - 2$ (Konishi 1996). Recent work in political economy often only considers a public good space of one dimension (Acemoglu and Robinson 2006). Note however, that if $\mathbb{D}$ is collegial, then $Cycle(W, H_{\mathbb{D}}(u)) = \Phi$and $C(W, H_{\mathbb{D}}(u)) \neq \Phi$. Such a rule can be called oligarchic, and this inference provides a theoretical basis for comparing democracy and oligarchy(Acemoglu 2008). Figure 3 showed the preference cones in a majority voting game with 3 agents and Nakamura number 3, so half openness fails in two dimensions.

Extending the equilibrium result of the Nakamura Theorem to higher dimension for a voting rule faces a difficulty caused by Saari's Theorem. We first define a *fine* topology on smooth utility functions (Hirsch 1976; Schofield 1999c; Schofield 2003).

Definition 5 Let $(U(W)^N, T_1)$ be the topological space of smooth utility profiles endowed with the C^1-topology.

In economic theory, the existence of isolated price equilibria can be shown to be "generic" in this topological space (Debreu 1970, 1976; Smale 1974a, b). In social choice no such equilibrium theorem holds. The difference is essentially because of the coalitional nature off social choice.

Saari Theorem

For any non-collegial $\mathbb{D}$, there exists an integer $w(\mathbb{D}) \geq \kappa(\mathbb{D}) - 1$ such that $dim(W) > \; w(\mathbb{D})$ implies that $C(W, H_{\mathbb{D}}(u)) = \Phi$ for all u in a dense subspace of $(U(W)^N, T_1)$ so $Cycle(W, H_{\mathbb{D}}(u)) \neq \Phi$ generically. □

This result was essentially proved by Saari (1997), building on earlier results by Plott (1967), McKelvey (1979), Schofield (1983a); Schofield (1983b), McKelvey and Schofield (1987) and Banks (1995). See Saari (1985a); Saari (1985b); Saari (2001a, b, 2008) for related analyses. Indeed, it can be shown that if $dim(W) > w(\mathbb{D}) + 1$ then $Cycle(W, H_{\mathbb{D}}(u))$ is generically dense (Schofield 1984c). The integer $w(\mathbb{D})$ can usually be computed explicitly from $\mathbb{D}$. For majority rule with n odd it is known that $w(\mathbb{D}) = 2$ while for n even, $w(\mathbb{D}) = 3$. Saari showed, for a $q - rule$ where any coalition of size at least $q(<n)$ belongs to the set of decisive coalitions, $\mathbb{D}_q$, that

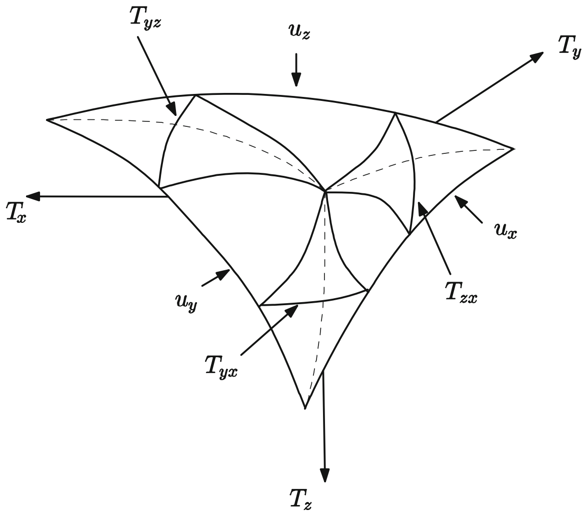
$$w(\mathbb{D}_q) = 2q - n + 1 + max\left\{\frac{4q - 3n - 1}{2(n-q)}, 0\right\}$$

Although the Saari Theorem formally applies only to voting rules, Schofield (2010) argues that it is applicable to any non-collegial social mechanism, say $H(u)$ and can be interpreted to imply that

$$Cycle(W, H(u)) \neq \Phi \;\text{ and } C(W, H(u)) = \Phi$$

is a generic phenomenon in coalitional systems. Because preference curves can wander anywhere in any open component of $Cycle(W, H(u))$, Schofield (1979) called this *chaos*. It is not so much the sensitive dependence on initial conditions, but the aspect of indeterminacy that is emphasized. To illustrate, Fig. 4 shows the preference cones in an prisoners' dilemma with three players. In this example the preference fields, $\{T_x, T_y, T_z,\}$ of the non-cooperative individuals belong to a different half space from the two person cooperative fields $\{T_{yx}, T_{yz}, T_{zx},\}$. Again half openness fails, so $Cycle(W, H(u)) \neq \Phi$, and we lack any result that would allow us to infer existence of a cooperative equilibrium. Indeed, cooperative and non-cooperative behavior are both possible. On the other hand, existence of a hegemon, as discussed in Sect. 2, is similar to existence of a collegium, suggesting that $Cycle(W, H(u))$ would be constrained in this case. Recent work by Wilson (2012) suggests that the nPD can be seen as a heuristic model for group selection as a basis for the evolution of *H.Sapiens.* As he points out there is a conflict between evolution based on selection based on individual fitness and group selection It follows from the logic of the nPD that the fitness of a selfish individual will exceed that of an altruistic individual, but the fitness of a group of altruists will exceed that of a group of selfish individuals. Consequently we should expect both selfishness and altruism to be selected for in the evolution of *H.Sapiens.* He suggests that our culture does indeed exhibit both traits. Indeed the argument about chaos underlying the nPD further suggests that the evolutionary path leading to *H.Sapiens* was indeed chaotic

Fig. 4 Failure of half-openness in the three person prisoners' dilemma



Richards (1993) has examined data on the distribution of power in the international system over the long run and presents evidence that it can be interpreted in terms of a chaotic trajectory. This suggests that the metaphor of the nPD in international affairs does characterise the ebb and flow of the system and the rise and decline of hegemony.

In the following sections I shall consider more general social processes in order to examine how $Cycle(W, H)$ may be subject to catastrophic change. The next section considers models of belief aggregation or truth-seeking in a society.

5 Beliefs and Condorcet's Jury Theorem

The Jury theorem formally only refers to a situation where there are just two alternatives $\{1, 0\}$, and alternative 1 is the "true" option. Further, for every individual, i, it is the case that the probability that i picks the truth is ρ_{i1}, which exceeds the probability, ρ_{i0}, that i does not pick the truth. We can assume that $\rho_{i1} + \rho_{i0} = 1$, so obviously $\rho_{i1} > \frac{1}{2}$. To simplify the proof, we can assume that ρ_{i1} is the same for every individual, thus $\rho_{i1} = \alpha > \frac{1}{2}$ for all i. We use $\chi_i (= 0 \text{ or } 1)$ to refer to the choice of individual i, and let $\chi = \Sigma_{i=1}^{n} \chi_i$ be the number of individuals who select the true option 1. We use Pr for the probability operator, and E for the expectation operator. In the case that the electoral size, n, is odd, then a majority, m, is defined to be $m = \frac{n+1}{2}$. In the case n is even, the majority is $m = \frac{n}{2} + 1$. The probability that a majority chooses the true option is then

$$\alpha_{maj}^{n} = \Pr[\chi \geq m].$$

The theorem assumes that voter choice is *pairwise independent*, so that $\Pr(\chi = j)$ is simply given by the binomial expression $\binom{n}{j}\alpha^j(1-\alpha)^{n-j}$.

A version of the theorem can be proved in the case that the probabilities $\{\rho_{i1} = \alpha_i\}$ differ but satisfy the requirement that $\frac{1}{n}\Sigma_{i=1}^n \alpha_i > \frac{1}{2}$. Versions of the theorem are valid when voter choices are not pairwise independent (Ladha and Miller 1996).

The Jury Theorem. If $1 > \alpha > \frac{1}{2}$, then $\alpha_{maj}^n \geq \alpha$, and $\alpha_{maj}^n \longrightarrow 1$ as $n \longrightarrow \infty$.
Proof Consider the case with n odd. Now

$$\Pr(\chi = j)\binom{n}{j}\alpha^j(1-\alpha)^{n-j} = \binom{n}{n-j}\alpha^j(1-\alpha)^{n-j} \tag{8}$$

Since $\alpha > \frac{1}{2}$ we see that $\alpha^{n-j}(1-\alpha)^j > \alpha^j(1-\alpha)^{n-j}$. Thus,

$$\Sigma_{j=0}^{m-1} j \Pr(\chi = n-j) > \Sigma_{j=0}^{m-1} j \Pr(\chi = j)$$

or

$$\Sigma_{k=m}^{n}(n-k)\Pr(\chi = k) > \Sigma_{k=0}^{m-1} k \Pr(\chi = k).$$

Thus

$$n\Sigma_{k=m}^{n} \Pr(\chi = k) > \Sigma_{k=0}^{m-1} k \Pr(\chi = k) + \Sigma_{k=m}^{n} k \Pr(\chi = k).$$

But

$$n\alpha_{maj} = n\Sigma_{k=m}^{n} \Pr(\chi = k)$$

and

$$E(\chi) = \Sigma_{k=0}^{n} k \Pr(\chi = k) = n\alpha.$$

Thus, $\alpha_{maj} > \alpha$ when n is odd.

The case with n even follows in similar fashion, taking $m = \frac{n}{2} + 1$, and using an equi-probable tie-breaking rule when $k = \frac{n}{2}$.This gives

$$\alpha_{maj} = \Sigma_{k=m}^{n} \Pr(\chi = k) + \frac{1}{2}\Pr(\chi = \frac{n}{2}).$$

For both n being even or odd, as $n \longrightarrow \infty$, the fraction of voters choosing option 1 approaches $\frac{1}{n}E(\chi) = \alpha > \frac{1}{2}$. Thus, in the limit, more than half the voters choose the true option. Hence the probability $\alpha_{maj}^n \longrightarrow 1$ as $n \longrightarrow \infty$. □

Laplace also wrote on the topic of the probability of an error in the judgement of a tribunal. He was concerned with the degree to which jurors would make just decisions in a situation of asymmetric costs, where finding an innocent party guilty was to be more feared than letting the guilty party go free. As he commented on the appropriate rule for a jury of twelve, "I think that in order to give a sufficient guarantee to innocence, one ought to demand at least a plurality of nine votes in twelve"Laplace (Laplace 1951[1814]:139). Schofield (1972a, b) considered a model derived from the jury theorem where uncertain citizens were concerned to choose an ethical rule which would minimize their disappointment over the likely outcomes, and showed that majority rule was indeed optimal in this sense.

Models of belief aggregation extend the Jury theorem by considering a situation where individuals receive signals, update their beliefs and make an aggregate choice on the basis of their posterior beliefs (Austen-Smith and Banks 1996). Models of this kind can be used as the basis for analysing correlated beliefs[18] and the creation of belief cascades (Easley and Kleinberg 2010).

Schofield (2002, 2006) has argued that Condorcet's Jury theorem provided the basis for Madison's argument in Federalist X (Madison 1999 [1787]) that the judgments of citizens in the extended Republic would enhance the "probability of a fit choice." However, Schofield's discussion suggests that belief cascades can also fracture the society in two opposed factions, as in the lead up to the Civil War in 1860.[19]

There has been a very extensive literature recently on cascades[20] but it is unclear from this literature whether cascades will be equilibrating or very volatile. In their formal analysis of cascades on a network of social connections, Golub and Jackson (2010) use the term *wise* if the process can attain the truth. In particular they note that if one agent in the network is highly connected, then untrue beliefs of this agent can steer the crowd away from the truth. The recent economic disaster has led to research on market behavior to see if the notion of cascades can be used to explain why markets can become volatile or even irrational in some sense (Acemoglu et al. 2010; Schweitzer et al. 2009). Indeed the literature that has developed in the last few years has dealt with the nature of herd instinct, the way markets respond to speculative behavior and the power law that characterizes market price movements.[21] The general idea is that the market can no longer be regarded as efficient. Indeed, as suggested by Ormerod (2001) the market may be fundamentally chaotic.

---

[18] Schofield (1972a, b), Ladha (1992, 1993, 1995, 1996), Ladha and Miller (1996).

[19] Sunstein (2006, 2011) also notes that belief aggregation can lead to a situation where subgroups in the society come to hold very disparate opinions.

[20] Gleick (1987), Buchanan (2001); Buchanan (2003), Gladwell (2002), Johnson (2002), Barabasi (2003); Barabasi (2011), Strogatz (2004), Watts (2002, 2003), Surowiecki (2005), Ball (2004), Christakis and Fowler (2011).

[21] See, for example, Mandelbrot and Hudson (2004), Shiller (2003); Shiller (2005), Taleb (2007), Barbera (2009), Cassidy (2009), Fox (2009).

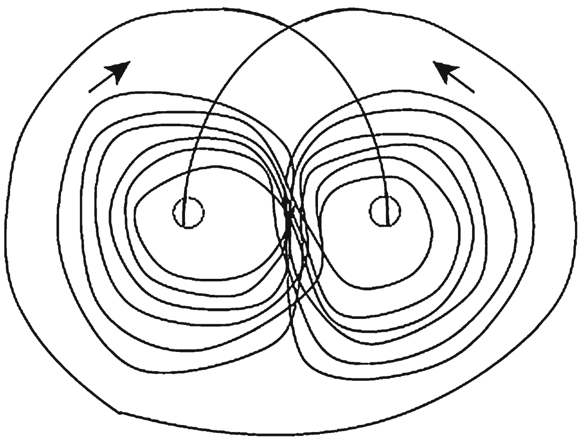


Fig. 5 The butterfly

"Empirical" chaos was probably first discovered by Lorenz (1962, 1963, 1993) in his efforts to numerically solve a system of equations representative of the behavior of weather.[22] A very simple version is the non-linear vector equation

$$\frac{dx}{dt} = \begin{bmatrix} dx_1 \\ dx_2 \\ dx_3 \end{bmatrix} = \begin{bmatrix} -a_1(x_1 - x_2) \\ -x_1 x_3 \ + \ a_2 x_1 - x_2 \\ x_1 x_2 - a_3 x_3 \end{bmatrix}$$

which is chaotic for certain ranges of the three constants, a_1, a_2, a_3.

The resulting "butterfly" portrait winds a number of times about the left hole (as in Fig. 5), then about the right hole,then the left, *etc.* Thus the "phase prortrait" of this dynamical system can be described by a sequence of winding numbers $(w_l^1, w_k^1, w_l^2, w_k^2,\ \mathit{etc.})$. Changing the constants a_1, a_2, a_3 slightly changes the winding numbers. Note that the picture in Fig. 5 is in three dimensions, The butterfly wings on left and right consist of infinitely many closed loops. The whole thing is called the Lorentz "strange attractor." A slight perturbation of this dynamic system changes the winding numbers and thus the qualitative nature of the process. Clearly this dynamic system is not structurally stable, in the sense used by Kauffman (1993). The metaphor of the butterfly gives us pause, since all dynamic systems whether models of climate, markets, voting processes or cascades may be indeterminate or chaotic.

[22] See also the earlier work by Richardson (1922) and by the team under von Neumann at the Institute for advanced Studies at Princeton using the computer known as ENIAC as described by Dyson (2012).

6 The Edge of Chaos

Recent work has attempted to avoid chaos by using the Brouwer fixed point theorem to seek existence of a *belief equilibrium* for a society N_τ of size $n_{\tau.}$ time τ. In this context we let

$$W_E = W_1 \times W_2 \ldots \times W_{n_{\tau+1.}} \times \Delta$$

be the economic product space, where W_i is the commodity space for citizen i and Δ is a price simplex. Let W_E be the economic space and $W_{\mathbb{D}}$ be a space of political goods, governed by a rule $\mathbb{D}$. At time τ, $W_\tau = W_E \times W_{\mathbb{D}}$ is the political economic space.

At τ, each individual, i, is described by a utility function $u_i : W_\tau \to \mathbb{R}$, so the population profile is given by $u : W_\tau \to \mathbb{R}^{n_\tau}$. Beliefs at τ about the future $\tau + 1$ are given by a stochastic rule, $\mathbb{Q}_\tau$, that transforms the agents' utilities from those at time τ to those at time $\tau + 1$. Thus $\mathbb{Q}_\tau$ generates a new profile for $N_{\tau+1}$ at $\tau + 1$ given by $\mathbb{Q}_\tau(u) = u' : W_{\tau+1} \to \mathbb{R}^{n_\tau+1}$. The utility and beliefs of i will depend on the various sociodemographic subgroups in the society $N_{\tau.}$ that i belongs to, as well as information about the current price vector in Δ.

Thus we obtain a transformation on the function space $[W_\tau \to \mathbb{R}^{\mathbf{n}_\tau}]$ given by

$$[W_\tau \to \mathbb{R}^{\mathbf{n}_\tau}] \to^{\mathbb{Q}_\tau}\to [W_\tau \to \mathbb{R}^{\mathbf{n}_{\tau+1}}] \to [W_\tau \to \mathbb{R}^{n_\tau}]$$

The second transformation here is projection onto the subspace $[W_\tau \to \mathbb{R}^{\mathbf{n}_\tau}]$ obtained by restricting to changes to the original population $N_{\tau.}$ and space.

A *dynamic belief equilibrium* at τ for $N_{\tau.}$ is fixed point of this transformation. Although the space $[W_\tau \to \mathbb{R}^{\mathbf{n}_\tau}]$ is infinite dimensional, if the domain and range of this transformation are restricted to *equicontinous* functions (Pugh 2002), then the domain and range will be compact. Penn (2009) shows that if the domain and range are convex then a generalized version of Brouwer's fixed point theorem can be applied to show existence of such a dynamic belief equilibrium. This notion of equilibrium was first suggested by Hahn (1973) who argued that equilibrium is located in the mind, not in behavior.

However, the choice theorem suggests that the validity of Penn's result will depend on how the model of social choice is constructed. For example Corcos et al. (2002) consider a formal model of the market, based on the reasoning behind Keynes's "beauty contest"(Keynes 1936). There are two coalitions of "bulls" and "bears". Individuals randomly sample opinion from the coalitions and use a *critical* cutoff-rule. For example if the individual is bullish and the sampled ratio of bears exceeds some proportion then the individual flips to bearish. The model is very like that of the Jury Theorem but instead of guaranteeing a good choice the model can generate chaotic flips between bullish and bearish markets, as well as fixed points or cyclic behavior, depending on the cut-off parameters. Taleb's argument (Taleb 1997) about black swan events can be applied to the recent transformation in societies in the Middle East and North Africa that resemble such a cascade (Taleb and Blyth 2011). As in the earlier episodes in Eastern Europe, it would seem plausible that the sudden onset of a cascade is due to a switch in a critical coalition.

The notion of "criticality" has spawned in enormous literature particularly in fields involving evolution, in biology, language and culture.[23] Bak and Sneppen (1993) refer to the self organized critical state as the

> "edge of chaos" since it separates a frozen inactive state from a "hot" disordered state.

> The mechanism of evolution in the critical state can be thought of as an exploratory search for better local fitness, which is rarely successful, but sometimes has enormous effect on the ecosystem

Flyvbjerg et al. (1993) go on to say

> species sit at local fitness maxima..and occasionally a species jumps to another maximum [in doing so it] may change the fitness landscapes of other species which depend on it. ..Consequently they immediately jump to new maxima. This may affect yet another species in a chain reaction, a *burst* of evolutionary activity.

This work was triggered by the earlier ideas on "punctuated equilibrium" by Eldredge and Gould (1972).[24] As Gould (2002:782) writes

> [T]hus, the pattern of punctuated equilibrium establishes species as effective individuals and potential Darwinian agents in the mechanisms of macroevolution.

There are a number of points to be made about these remarks. First of all, the fitness model can be viewed as essentially a game, based on a search for a critical (or local) Nash equilibrium.[25] When Dawkins (1976) wrote of the "selfish gene" he seemed to imply that evolution could be structured in the form of a non-cooperative game But Jablonka and Lamb (2006:38) observe

> For Gould, the central focus of evolutionary studies has to be organisms, groups, and species. For Dawkins, it has to be the gene, the unit of hereditary.

This suggests that "cooperation" is at the center of evolution. Following Gould, the critical Nash equilibria at the level of species can be destroyed by a "catastrophe"(Zeeman 1977). Finding the new evolutionary trajectory may require joint changes in a *coalition* of the species. Indeed there may be numerous different co-adaptive species.

Second, evolutionary transformations at all levels are largely the result of new configurations of coalitions of genes. For Margulis and Sagan (2002:12, 20),

> the major source of inherited variation is not random mutation. Rather the important transmitted variation that leads to evolutionary novelty comes from the acquisition of genomes. Entire sets of genes.. are acquired and incorporated by others.

---

[23] See for example Cavallli-Sforza and Feldman (1981), Bowles et al. (2003).

[24] See also Eldredge (1976), Gould (1976).

[25] See Maynard Smith (1982) for the game theoretic notion of evolutionary stable strategy.

> We must begin to think of organisms as communities, as collectives. And communities are ecological entities.

At the level of the gene, evolutionary change may requires new genomic structures that again are coalitions of genes.

The point to be emphasized is that the evolution of a species involves bifurcation, a splitting of the pathway. We can refer to the bifurcation as a *catastrophe* or a *singularity*. The portal or door to the singularity may well be characterized by chaos or uncertainty, since the path can veer off in many possible directions, as suggested by the bifurcating cones in Figs. 3 and 4. At every level that we consider, the bifurcations of the evolutionary trajectory seem to be locally characterized by chaotic domains. I suggest that these domains are the result of different coalitional possibilities. The fact that the trajectories can become indeterminate suggests that this may enhance the exploration of the fitness landscape.

A more general remark concerns the role of climate change. Climate has exhibited chaotic or catastrophic behavior in the past.[26] There is good reason to believe that human evolution over the last million years can only be understood in terms of "bursts" of sudden transformations (Nowak 2011) and that language and culture co-evolve through group or coalition selection (Cavallli-Sforza and Feldman 1981). Calvin (2003) suggests that our braininess was cause and effect of the rapid exploration of the fitness landscape in response to climatic forcing.

7 Structural Stability

Instead of looking for chaos, we can use the method of Sect. 3 and consider a generalized preference field $H : W_\tau \twoheadrightarrow TW_\tau$ to model the flow generated by H on a state space W_τ. We could consider a general evolutionary process, but for the moment we will focus on the a field H associated with a society of individuals. By extending the flow over some time interval $\Upsilon = [\tau, \tau + 1]$, we could then examine whether it appears to be structurally stable.

In the above, we have used the term *structural stability* for the property that the qualitative features of the flow are not changed by small perturbations in the underlying parameters. The term *chreod* was used by Thom ([1975], 1994) in the context of evolutionary or biological processes to describe such a dynamical system that returns to a steady trajectory. This term is derived from the Greek word for "necessary" and the word for "pathway". A *social chreod* is therefore a structurally stable path through time, where the state space W_Υ for human society now involves not only characteristics, such as factor endowments and prices, p, in

[26] Indeed as I understand the dynamical models, the chaotic episodes are due to the complex interactions of dynamical processes in the oceans, on the land, in weather, and in the heavens. These are very like interlinked *coalitions* of non-gradient vector fields.

W_E, but also the beliefs and thus the preferences of individuals, particularly as regards the risk postures that are embedded in their preferences.

One advantage of such a modeling exercise is that we do not need not need to use Penn's equilibrium theorem to model the transformation $\mathbb{Q}_\tau(u) = u'$. Instead, the society N_τ is characterised at τ by a family $\{du_i : W_\tau \to T^*W_\tau : i \in N_\tau\}$ of normalized direction gradients. Here $du_i(x,p) : TW_\tau \to \mathbb{R}$ specifies the local change in utility for i at a point x, and price p, when composed with a vector $v \in TW_\tau$. Just as in the model for the choice theorem, we can then define, for any coalition $M \subset N_\tau$, the preference field $H_M(x) \subset T_xW_\tau$ at $x \in W_\tau$, consisting of vectors in T_xW_τ that are both preferred and feasible for the coalition at x. Taking $H = \cup H_M \cup H_\Delta$, where H_Δ is a field on the price simplex, gives the generalised field $H : W_\Upsilon \twoheadrightarrow TW_\Upsilon$ in the tangent bundle TW_Υ above W_Υ.

Such a flow could induce changes in beliefs about the future sufficient to cause discontinuities or bifurcations in the preference field, H. On the other hand, there may be conditions under which the field H is half open. If indeed, H is half open on TW_Υ, then there will exist a *local social direction gradient*, $d : W_\Upsilon \to T^*W_\Upsilon$ with the property that $d((x)(v) > 0$ for every $v \in H(x)$ at every $x \in W_\Upsilon$(where $H(x) \neq \Phi$).

To put together the political and economic models, we now consider, for each individual i at time τ and state $x \in W_\tau$ a stochastic preference cone $H_i(x) \subset T_xW$. That is, let

$$\{du_i : W \to T^*W_\tau : i \in N_\tau\}$$

define the set of utility direction gradients for the society, N_τ, at time τ. Now let $H_i(x)$ be a generalized cone or probability distribution over the set of vectors

$$H_i(x) = \{v \in T_xW : du_i(x)(v) > 0\}$$

The *acyclicity* or *half open* condition at x is the property that

$$[\cup H_i(x)]^* = \cap_i[H_i(x)]^* \neq \Phi\}.$$

If this condition is satisfied then we can chose a vector $d(x) \in \cap[H_i(x)]^*$. Extending d over a neighborhood then gives a social preference path $c_{N,\tau} : [-1. + 1] \twoheadrightarrow W_\Upsilon$ in the time interval $\Upsilon = [\tau, \tau + 1]$.

The aggregate social path in W_Υ will of course involve changes in the price vector, along the lines discussed above. The jumps occasioned by a policy switch at time, τ, may very well induce changes in the induced social preference cone $H_{\tau+1}$ at time $\tau + 1$. However, if the dual field $H^*_{\tau+1}$ is non empty and lower hemi-continuous, then by Michael's selection theorem (Michael 1956) it would allow a continuous selection, namely a social direction gradient d defined in a neighbor-hood of x.

In line with the previous results however, we can only guarantee that the path is well defined if the cone $H_{\tau+1}$ has a non empty dual $H^*_{\tau+1}$ everywhere. This dual field can be regarded either as a set of social direction gradients, or as a set of "social shadow prices." Since the dual may be empty at some time τ, we can instead modify the preference field, by defining the *heart* of a preference field.

Definition 6: The Heart

(i) If W is a topological space, then $x \in W$ is *locally covered* (under the preference correspondence Q) iff for any neighborhood Y of x in W, there exists $y \in Y$ such that (a) $y \in Q(x)$, and (b) there exists a neighborhood Y' of y, with $Y' \subseteq Y$ such that $Y' \cap Q(y) \subset Q(x)$.In this case write $y \in \mathbf{Q}(x)$.
(ii) The *heart* of Q, written $\mathcal{H}(Q)$, is the set of locally uncovered points in W, namely $\mathcal{H}(Q) = \{x : \mathbf{Q}(x) = \Phi\}$.

That is, we construct the preference relation by following the paths generated by the original field H. We write $\mathbb{H}(x)$ for the set of points reachable from x by such a path and then eliminate possible paths from x if there is another point y near x such that y is reachable from x, and all points reachable from y are also reachable from x. By this method we restrict the paths that are allowed, and obtain a new field $\mathbf{H}$ and define $\mathcal{H}(\mathbf{H}) = \{x \in W : \mathbf{H}(\mathbf{x}) = \Phi\} \subset W$ to be the heart induced from H. Under fairly general conditions, the heart will be non-empty and will belong to the Pareto set of the utility profile u of the society.[27] Regarded as a correspondence from the set of parameters of the model into $W, \mathcal{H}$ is lower hemicontinuous and its image at $\mathbf{H}$, namely $\mathcal{H}(\mathbf{H})$, is star-shaped. Schofield (1999a, b) suggests that actual outcomes can be derived as a selection from the heart correspondence.[28]

Example 2 To illustrate the heart, Fig. 6 gives a simple artificial example where the utility profile, u, is one where society has "Euclidean" preferences, based on distance, and the ideal points are uniformly distributed on the boundary of the equilateral triangle. Under majority rule, $\mathbb{D}$, the heart $\mathcal{H}(\mathbf{H}_{\mathbb{D}}(\mathbf{u}))$, is the star-shaped figure inside the equilateral triangle (the Pareto set), and contains the "yolk"(McKelvey 1986). The heart is generated by an infinite family of "median lines," such as $\{M_{1,}M_{2,}..\}$. The shape of the heart reflects the asymmetry of the distribution. Inside the heart, voting trajectories can wander anywhere. Indeed, a result of Tataru (1999) shows that the reachability correspondence $\mathbb{H}_{\mathbb{D}}(u)$, for the majority–rule, $\mathbb{D}$, expands linearly inside the heart. Outside the heart the dual cones intersect, so any trajectory starting outside the heart converges to the heart. Thus the heart is an "attractor."

With sufficient change in the parameters the heart collapses. In Fig. 7 the population is uniformly distributed on the boundary of the more symmetric pentagon, so the heart is very small.[29] If the ideal points were distributed symmetrically on the circle or the sphere (in higher dimensions) then the Plott (1969)

---

27 This can be shown adapting the proof technique of Banks et al. (2002, 2006).

28 This depends on the extension of Michael's selection theorem by Mas-Colell (1979).

29 It is bounded by median arcs such as (M_1, M_2).

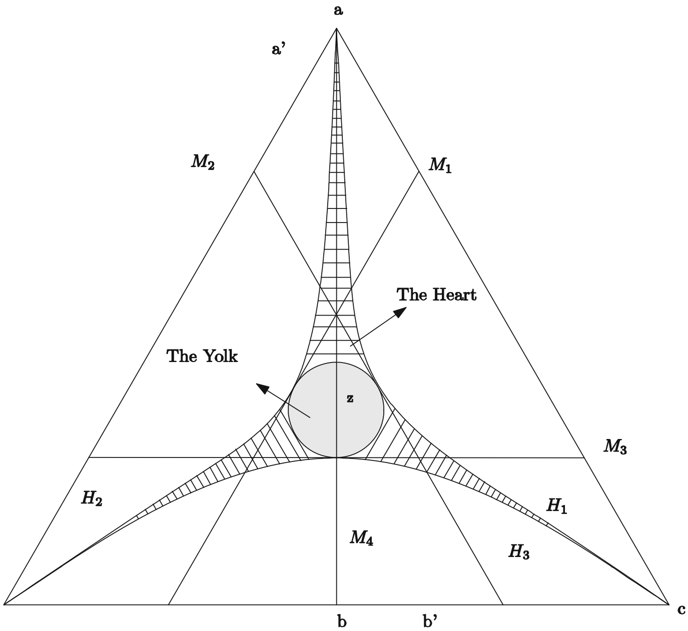
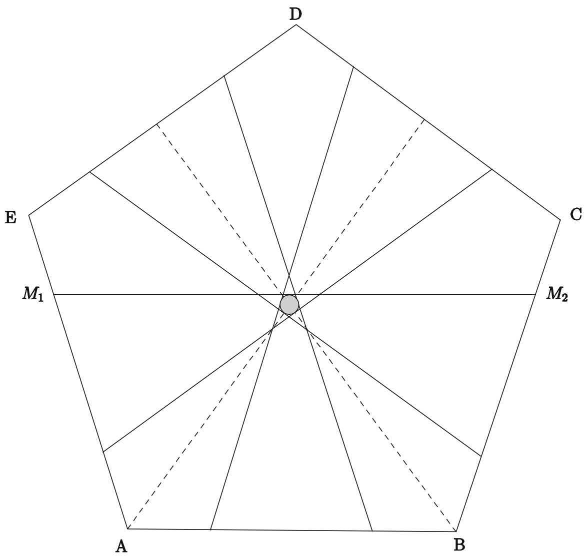


Fig. 6 The heart with a uniform electorate on the triangle

Fig. 7 The heart with a uniform electorate on the pentagon



symmetry conditions would be satisfied, and heart would coincide with a point core or equilibrium.[30]

Now consider a major general dynamical process of the kind considered above. If indeed the non-empty dual or the acyclicity property is satisfied then the dual field will permit a local social direction gradient, which can be followed in the direction of the Pareto set. It is still possible however that, at some time τ, the dual condition on the social preference cones will fail, in which case local chaos is possible in the neighborhood of a bifurcation. In some cases the trajectory can be considered to be treelike, but at each bifurcation there will be a domain of local indeterminate behavior. Under some conditions this domain may collapse in which case the trajectory will continue after the local chaos, but in an unpredictable fashion.

We can use these figures as a heuristic device to think about an evolutionary process. If Fig. 6 represents the evolutionary landscape, then no agent (whether species or voter) can guarantee an optimum position. Bak (1996:122) calls this the "Red Queen effect" and describes simple models that generate probability distributions obeying a power law of the form $p(x) = x^{-\beta}$, where the exponent, β, is often found to lie in the range $[1, 2]$. In such models the log-log figure gives a straight line. [Power laws were first proposed by Pareto (1935), to characterize wealth and by Zipf (1965), to describe city size and word frequency.] In Fig. 6 it is obvious that most events lie near the center, the yolk, while extreme (low probability) events lie near the vertices of the heart. Tataru's result on the reachability correspondence for voting suggest that the jumps inside the heart do not follow a Gaussian or Poisson distribution but rather a power law. Applying this to evolution, we can infer that when the trajectory approaches one of the vertices then agents on the opposed face may become extinct. Thus we would obtain a power law in extinctions. In contrast, as the distribution becomes symmetric then the heart collapses to a point equilibrium, which can be regarded as an evolutionary stable state, as proposed by Maynard Smith (1982). The most important characteristic of the heart correspondence is that it is lower-hemicontinuous. What this means is that at some parameter values a slight change in these values causes the heart to explode from a small "equilibrium" to a much larger domain governed by unpredictability. I shall call this event *inflation*.[31] This seems to be precisely what Gould had in mind for evolutionary processes, and what Taleb and others have observed as the sudden onset of volatility in markets.

In writing about "virtual history," Ferguson (1997) commented that it was based on

> calculations about the relative probability of plausible outcomes in a chaotic world.

[30] See also Arrow (1969) and later analyses of such games in Schofield and Tovey (1992).

[31] I use this as a metaphor derived from the notion of inflation in cosmology (Penrose 2011). If we can use the term entropy to characterise the distribution of events in the heart, then entropy increases dramatically at the inflationary point.

It is this idea of chaotic bifurcation in a social trajectory that provides the motivation for the heart. It seems that the uncertainty implicit in this extremely complex dynamical game provides a justification for the assertion by Popper (1959) that prediction is impossible in the social sciences. Barabasi (2011) argues, in contrast, that our ability to process data will allow us in the future to model society in a deterministic fashion. However, even attempts to analyze climate change itself has shown how difficult it is to model inter-related processes with positive feedback mechanisms (Edwards 2010).

If my interpretation of the models presented here is correct, then uncertainty is a fundamental fact of coalitional forces in society. In particular the uncertainty involves the critical transition of inflation from what appears to be an equilibrium situation to a situation where the trajectory becomes highly volatile. This seems to be the lesson to be drawn from Arrow's Impossibility theorem.

8 Conclusion

I do not attempt to explore the full ramifications of this model here. In principle this model can be adapted to situations where both preferences and beliefs are aggregated. More generally I suggest it can be used where there are multiple interacting dynamical systems, such as in climate change. Note however, that cycling can still occur within the heart. Intuitively in the society or an economy, the more disparate the beliefs the more extensive will be the heart and thus the degree of cycling. The greater the concentration of power or the centrality of the underlying social network, the smaller the heart and the less the importance of cycling. As this essay has pointed out, all the equilibrium theorems in economics or social choice theory are based on assumptions of convexity and/or acyclicity and thus essentially on the condition that the dual cones intersect at *every* possible state. This would seem to be a non-generic property. Since political and economic behavior are both coalition based, there is little reason to expect equilibrium as it is usually understood. The notion of the heart is more general that of equilibrium and can be used in situations where there are coalitional cycles. Note, however, that at an inflationary event where an equilibrium evaporates, the developmental path will become unpredictable.

Applications of the model presented here to the study of US elections (Miller and Schofield 2003, 2008) suggest that the extreme unpredictability indicated by the voting theorem depends on details of the political system. For example Fig. 8 gives an estimate of the voter locations in the 2008 presidential election (Schofield and Schnidman 2011). The voter distribution is somewhat skew symmetric so the heart is fairly small in the two dimensional space. Over time there can be cyclic behavior involving both " circumferential" and "radial" transformations as the heart shifts in response to changes in beliefs. Cycles in the market may well induce cyclic political behavior. In a polity run by an autocrat, such as Syria, the heart can be identified with the position of the autocrat. The removal of the autocrat, as in

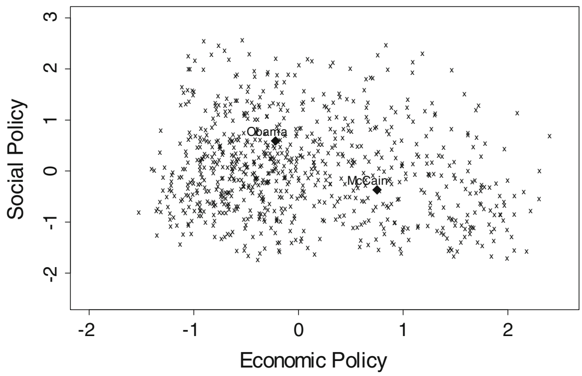


Fig. 8 Distribution of voter ideal points and candidate positions in the 2008 presidential election

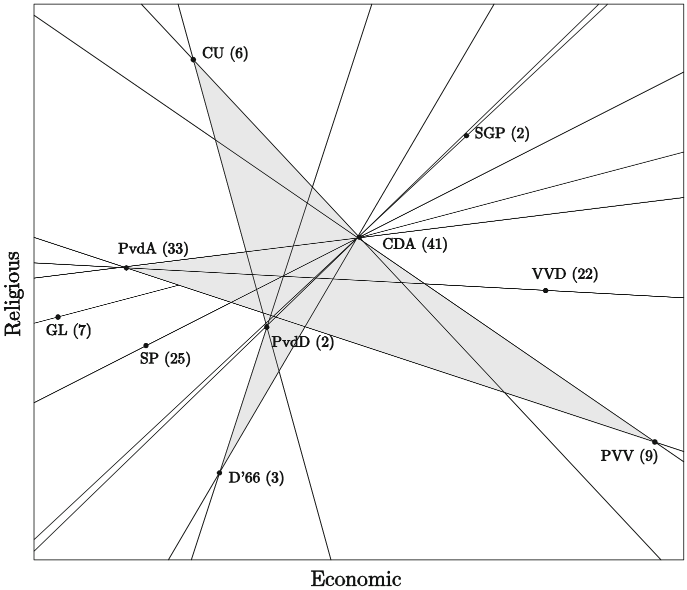


Fig. 9 Party positions in the Netherlands in 2006

Egypt, creates what we have called inflation, so that the likely future becomes very unpredictable. In European democracies, based on a proportional electoral system, the existence of small and possibly extreme parties implies a large political heart, as shown in Fig. 9, for the case of the Netherlands. Again, this suggests a degree of instability in government. In the international arena, the policy differences between the U.S., Europe and the BRICS can be expected to generate a very large heart, unless there is a degree of hegemonic leadership of the kind discussed in Sect. 2.

As regards fundamental choices over how to deal with climate change, it would seem necessary to attain sufficient agreement in beliefs so as to be able to define a local social direction gradient, obtained as a result of the non-empty intersection of the duals of the belief cones of the society. It appears very unlikely that this dual condition would hold everywhere, but political leadership could bring about the situation where there was some degree of consensus associated with a relatively constrained political heart.

Parfit's remarks on climate change are worth quoting here:

> What matters most is that we rich people give up some of our luxeries, ceasing to overheat the Earth's atmosphere, and taking care of this planet in other ways, so that it continues to support intelligent life. If we are the only rational animals in the Universe, it matters even more whether we shall have descendants during the billions of years in which that would be possible. Some of our descendants might live lives and create worlds that, though failing to justify past suffering, would give us all, including those who suffered, reason to be glad that the Universe exists. (Parfit 2011:419)

References

Acemoglu, D. (2008). Oligarchic versus democratic societies. *Journal of the European Economic Association*, *6*, 1–44.

Acemoglu, D., & Robinson, J. (2006). *Economic origins of dictatorship and democracy*. Cambridge: Cambridge University Press.

Acemoglu, D., & Robinson, J. (2008). Persistence of power, elites, and institutions. *American Economic Review*, *98*, 267–293.

Acemoglu, D., & Robinson, J. (2011). *Why nations fail*. London: Profile Books.

Acemoglu, D., Johnson, S., Robinson, J., & Yared, P. (2009). Reevaluating the modernization hypothesis. *Journal of Monetary Economics*, *56*, 1043–1058.

Acemoglu, D., Ozdaglar, A., & Tahbaz-Salehi, A. (2010). *Cascades in networks and aggregate volatility*. NBER Working Paper # 16516.

Akerlof, G. A., & Shiller, R. J. (2009). *Animal spirits*. Princeton: Princeton University Press.

Aliprantis, C., & Brown, D. J. (1983). Equilibria in markets with a Riesz space of commodities. *Journal of Mathematical Economics*, *11*, 189–207.

Arrow, K. J. (1951). *Social choice and individual values*. New Haven: Yale University Press.

Arrow, K. J. (1969). Tullock and an existence theorem. *Public Choice*, *6*, 105–111.

Arrow, K. (1986). Rationality of self and others in an economic system. *Journal of Business*, *59*, S385–S399.

Arrow, K., & Debreu, G. (1954). Existence of an equilibrium for a competitive economy. *Econometrica*, *22*, 265–90.

Austen-Smith, D., & Banks, J. (1996). Information aggregation, rationality, and the Condorcet Jury Theorem. *American Political Science Review*, *90*, 34–45.

Axelrod, R. (1981). The emergence of cooperation among egoists. *American Political Science Review*, *75*, 306–318.

Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, *211*, 1390–1396.

Bak, P. (1996). *How nature works: The science of self oerganized criticality*. Berlin: Springer.

Bak, P., & Sneppen, K. (1993). Punctuated equilibrium and criticality in a simple model of evolution. *Physical Review Letters*, *71*(24), 4083–4086.

Ball, P. (2004). *Critical mass*. New York: Ferrar, Strauss and Giroux.
Banks, J. S. (1995). Singularity theory and core existence in the spatial model. *Journal of Mathematical Economics*, *24*, 523–536.
Banks, J., Duggan, J., & Le Breton, M. (2002). Bounds for mixed strategy equilibria and the spatial model of elections. *Journal of Economic Theory*, *103*, 88–105.
Banks, J., Duggan, J., & Le Breton, M. (2006). Social choice and electoral competition in the general spatial model. *Journal of Economic Theory*, *126*, 194–234.
Barabasi, A.-L. (2003). *Linked*. New York: Plume.
Barabasi, A.-L. (2011). *Bursts*. New York: Dutton.
Barbera, R. (2009). *The cost of capitalism: Understanding market mayhem*. New York: McGraw Hill.
Bergstrom, T. (1975). *The existence of maximal elements and equilibria in the absence of transitivity*. Typescript: University of Michigan.
Bergstrom, T. (1992). When non-transitive relations take maxima and competitive equilibrium can't be beat. In R. Riezman & W. Neuefeind (Eds.), *Economic theory and international trade*. Berlin: Springer.
Bikhchandani, S., Hirschleifer, D., & Welsh, I. (1992). A theory of fads, fashion, custom, and cultural change as information cascades. *Journal of Political Economy*, *100*, 992–1026.
Binmore, K. (2005). *Natural justice*. Oxford: Oxford University Press.
Binmore, K. (2009). *Rational decisions*. Princeton, NJ: Princeton University Press.
Bowles, S., et al. (2003). The co-evolution of individual behaviors and socal institutions. *Journal of Theoretical Biology*, *223*, 135–147.
Boyd, J., & Richerson, P. J. (2005). *The origin and evolution of culture*. Oxford: Oxford University Press.
Brown, R. (1971). *The Lefschetz fixed point theorem*. Glenview, IL: Scott and Foreman.
Brouwer, L. E. J. (1912). Uber abbildung von mannigfaltikeiten. *Mathematische Annalen*, *71*, 97–115.
Buchanan, M. (2001). *Ubiquity*. New York: Crown.
Buchanan, M. (2003). *Nexus*. New York: Norton.
Burkhart, J. M., Hrdy, S. B., & van Schaik, C. P. (2009). Cooperative breeding and human cognitive evolution. *Evolutionary Anthropology*, *18*, 175–186.
Calvin, W. H. (2003). *The ascent of mind*. New York: Bantam.
Carothers, T. (2002). The end of the transition paradigm. *Journal of Democracy*, *13*, 5–21.
Cassidy, J. (2009). *How markets fail: The logic of economic calamities*. New York: Farrar, Strauss and Giroux.
Cavallli-Sforza, L., & Feldman, M. (1981). *Cultural transmission and evolution*. Princeton, NJ: Princeton University Press.
Chichilnisky, G. (1992). Social diversity, arbitrage, and gains from trade: A unified perspective on resource allocation. *American Economic Review*, *84*, 427–434.
Chichilnisky, G. (1993). Intersecting families of sets and the topology of cones in economics. *Bulletin of the American Mathematical Society*, *29*, 189–207.
Chichilnisky, G. (1995). Limited arbitrage is necessary and sufficient for the existence of a competitive equilibrium with or without short sales. *Economic Theory*, *5*, 79–107.
Chichilnisky, G. (1996a). Markets and games: A simple equivalence among the core, equilibrium and limited arbitrage. *Metroeconomica*, *47*, 266–280.
Chichilnisky, G. (1996b). An axiomatic approach to sustainable development. *Social Choice and Welfare*, *13*, 231–257.
Chichilnisky, G. (1997a). A topological invariant for competitive markets. *Journal of Mathematical Economics*, *28*, 445–469.
Chichilnisky, G. (1997b). Limited arbitrage is necessary and sufficient for the existence of a equilibrium. *Journal of Mathematical Economics*, *28*, 470–479.
Chichilnisky, G. (1997c). Market arbitrage, social choice and the core. *Social Choice and Welfare*, *14*, 161–198.

Chichilnisky, G. (2000). An axiomatic approach to choice under uncertainty with catastrophic risks. *Resource Energy Economics*, *22*, 221–231.
Chichilnisky, G. (2009a). The topology of fear. *Journal of Mathematical Economics*, *45*, 807–816.
Chichilnisky, G. (2009b). *Avoiding extinction: Equal treatment of the present and the future*. Working Paper, Columbia University.
Chichilnisky, G. (2010). The foundations of statistics with black swans. *Mathematical Social Sciences*, *59*, 184–192.
Chichilnisky, G., & Eisenberger, P. (2010). Asteroids: Assessing catastrophic risks. *Journal of Probability and Statistics* Article ID 954750.
Chichilnisky, G. (2011a) Catastrophic risks with finite or infinite states. *International Journal of Ecological Economics and Statistics*, *23*, F11.
Chichilnisky, G. (2011b). Sustainable markets with short sales. *Economic Theory* (in press).
Christakis, N., & Fowler, J. H. (2011). *Connected*. New York: Back Bay.
Collier, P. (2009). *Wars, guns and votes*. New York: Harper.
Collier, P. (2010). *The plundered planet*. Oxford: Oxford University Press.
Condorcet, N. (1994 [1785]) Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. Paris: Imprimerie Royale. Translated in part in: McLean I, Hewitt F, Condorcet: Foundations of social choice and political theory. London: Edward Elgar.
Corcos, A., et al. (2002). Imitation and contrarian behavior: Hyperbolic bubbles, crashes and chaos. *Quantitative Finance*, *2*, 264–281.
Dasgupta, P. (2005). Three conceptions of intergenerational Justice. In H. Lillehammer & D. H. Mellor (Eds.), *Ramsey's legacy*. Oxford: Clarendon Press.
Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.
Debreu, G. (1970). Economies with a finite number of equilibria. *Econometrica*, *38*, 387–392.
Debreu, G. (1976). The application to economics of differential topology and global analysis: Regular differentiable economies. *American Economic Review*, *66*, 280–287.
Deutscher, G. (2006). *The unfolding of language*. New York: Holt.
Dyson, G. (2012). *Turing's Cathedral*. New York: Pantheon.
Easley, D., & Kleinberg, J. (2010). *Networks, crowds and markets*. Cambridge: Cambridge University Press.
Edwards, P. N. (2010). *A vast machine: Computer models, climate data, and the politics of global warming*. Cambridge, MA: MIT Press.
Eldredge, N. (1976). Differential evolutionary rates. *Paleobiology*, *2*, 174–177.
Eldredge, N., & Gould, S. J. (1972). Punctuated equilibrium. In T. Schopf (Ed.), *Models of paleobiology*. New York: Norton.
Fan, K. (1961). A generalization of Tychonoff's fixed point theorem. *Mathematische Annalen*, *42*, 305–310.
Ferguson, N. (1997). Introduction. In N. Ferguson (Ed.), *Virtual history*. London: Picador.
Ferguson, N. (2002). *Empire: The rise and demise of the British world order*. London: Penguin Books.
Ferguson, N. (2011). *Civilization*. London: Penguin.
Flyvbjerg, H., Sneppen, K., & Bak, P. (1993). A mean field theory for a simple model of evolution. *Physical Review Letters*, *71*, 4087–4090.
Fox, J. (2009). *The myth of the rational market*. New York: Harper.
Fukuyama, F. (2011). *The origins of political order*. New York: Ferrar, Strauss and Giroux.
Gazzaniga, M.S. (2008). *Human*. New York: Harper.
Gintis, H. (2000). Strong reciprocity and human sociality. *Journal of Theoretical Biology*, *206*, 169–179.
Gladwell, M. (2002). *The tipping point*. New York: Back Bay.
Gleick, J. (1987). *Chaos: Making a new science*. New York: Viking.

Golub, B., & Jackson, M. (2010). Naive learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, *2*, 112–149.

Gould, S. J. (1976). The gnemoic metronome as a null hypothesis. *Paleobiology*, *2*, 177–179.

Gould, S. J. (2002). *The structure of evolutionary theory*. Cambridge, MA: Belknap Press of Harvard University Press.

Hahn, F. (1973). *On the notion of equilibrium in economics*. Cambridge: Cambridge University Press.

Hamilton, W. (1964). The genetical evolution of social behavior I and II. *Journal of Theoretical Biology*, *7*, 1–52.

Hamilton, W. (1970). Selfish and spiteful behavior in an evolutionary model. *Nature*, *228*, 1218–1220.

Hardin, G. (1968 [1973]). The tragedy of the commons. In H. E. Daly (Ed.), *Towards a steady state economy*. San Francisco: Freeman.

Hardin, R. (1971). Collective action as an agreeable prisons' dilemma. *Behavioral Sciences*, *16*, 472–481.

Hardin, R. (1982). *Collective action*. Baltimore: Johns Hopkins University Press.

Henrich, J., et al. (2004). *Foundations of human sociality*. Oxford: Oxford University Press.

Henrich, J., et al. (2005).'Economic man' in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences*, *28*, 795–855.

Hirsch, M. (1976). *Differential topology*. Berlin: Springer.

Hobbes, T. (2009 [1651]). In J. C. A. Gaskin (Ed.), *Leviathan; or the matter, forme, and power of a common-wealth, ecclesiastical and civil*. Oxford: Oxford University Press.

Hrdy, S. B. (2011). *Mothers and others: The evolutionary origins of mutual understanding*. Cambridge, MA: Harvard University Press.

Jablonka, E., & Lamb, M. J. (2006). *Evolution in four dimensions: Genetic, epigenetic, behavioral, and symbolic variation in the history of life*. Cambridge, MA: MIT Press.

Johnson, S. (2002). *Emergence*. New York: Scribner.

Karklins, R., & Petersen, R. (1993). Decision calculus of protestors and regime change: Eastern Europe 1989. *Journal of Politics*, *55*, 588–614.

Kauffman, S. (1993). *The origins of order*. Oxford: Oxford University Press.

Keohane, R. (1984). *After hegemony*. Princeton, NJ: Princeton University Press.

Keohane, R., & Nye, R. (1977). *Power and interdependence*. New York: Little Brown.

Keynes, J. M. (1936). *The general theory of employment, interest and money*. London: Macmillan.

Kindleberger, C. (1973). *The world in depression 1929–1939*. Berkeley, CA: University of California Press.

Knaster, B., Kuratowski, K., & Mazurkiewicz, S. (1929). Ein beweis des fixpunktsatzes fur n-dimensionale simplexe. *Fundamenta Mathematicae*, *14*, 132–137.

Konishi, H. (1996). Equilibrium in abstract political economies: With an application to a public good economy with voting. *Social Choice and Welfare*, *13*, 43–50.

Kramer, G. H. (1973). On a class of equilibrium conditions for majority rule. *Econometrica*, *41*, 285–297.

Kreps, D. M., et al. (1982). Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory*, *27*, 245–252.

Ladha, K. (1992). Condorcet's jury theorem, free speech and correlated votes. *American Journal of Political Science*, *36*, 617–674.

Ladha, K. (1993). Condorcet's jury theorem in the light of de Finetti's theorem: Majority rule with correlated votes. *Social Choice and Welfare*, *10*, 69–86.

Ladha, K., & Miller, G. (1996). Political discourse, factions and the general will: correlated voting and Condorcet's jury theorem. In N. Schofield (Ed.), *Collective decision making*. Boston: Kluwer.

Laplace, P. S. (1951 [1814]). Essai philosophique sur les probabilités. Paris: Gauthiers-Villars. A philosophical essay on probabilities. (trans. Truscott F, Emory F) New York: Dover.

Li, T. Y., & Yorke, J. A. (1975). Period three implies chaos. *American Mathematical Monthly*, *82*, 985–992.

Lohmann, S. (1994). The dynamics of Information cascades. *World Politics*, *47*, 42–101.

Lorenz, E. N. (1962). The statistical prediction of solutions of dynamical equations. *Proceedings International Symposium on Numerical Weather Prediction*, Tokyo.

Lorenz, E. N. (1963). Deterministic non periodic flow. *Journal of the Atmospheric Sciences, 130*, 141.

Lorenz, E. N. (1993). *The essence of chaos*. Seattle: University of Washington Press.

Madison, J. (1999). Federalist X. In J. Rakove (Ed.), *Madison: Writings*. New York: Library Classics.

Mandelbrot, B., & Hudson, R. (2004). *The (mis)behavior of markets*. New York: Perseus.

Margolis, H. (1982). *Selfishness, altruism and rationality*. Cambridge: Cambridge University Press.

Margulis, L., & Sagan, D. (2002). *Acquiring genomes*. New York: Basic.

Mas-Colell, A. (1979). A selection theorem for open graph correspondences with star-shaped values. *Journal of Mathematical Analysis and Applications*, *68*, 273–275.

Maynard Smith, J. (1982). *Evolution and the theory of games*. Cambridge: Cambridge University Press.

McKelvey, R. D. (1976). Intransitivities in multidimensional voting models and some implications for agenda control. *Journal of Economic Theory*, *12*, 472–482.

McKelvey, R. D. (1979). General conditions for global intransitivities in formal voting models. *Econometrica*, *47*, 1085–1112.

McKelvey, R. D. (1986). Covering, dominance and institution free properties of social choice. *American Journal of Political Science*, *30*, 283–314.

McKelvey, R. D., & Schofield, N. (1987). Generalized symmetry conditions at a core point. *Econometrica*, *55*, 923–933.

McWhorter, J. (2001). *The power of Babel*. New York: Holt.

Merton, R. C. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management*, *4*, 141–183.

Michael, E. (1956). Continuous selections I. *Annals of Mathematics*, *63*, 361–382.

Miller, G., & Schofield, N. (2003). Activists and partisan realignment in the U.S. *American Political Science Review*, *97*, 245–260.

Miller, G., & Schofield, N. (2008). The transformation of the Republican and Democratic party coalitions in the United States. *Perspectives on Politics*, *6*, 433–450.

Minsky, H. (1975). *John Maynard Keynes*. New York: Columbia University Press.

Minsky, H. (1986). *Stabilizing an unstable economy*. New Haven: Yale University Press.

Nakamura, K. (1979). The vetoers in a simple game with ordinal preference. *International Journal of Game Theory*, *8*, 55–61.

Mokyr, J. (2005). The intellectual origins of modern economic growth. *Journal of Economic History*, *65*, 285–351.

Mokyr, J. (2010). *The enlightened economy: An economic history of Britain 1700–1850*. New Haven: Yale University Press.

Mokyr, J., & Nye, V. C. (2007). Distributional coalitions, the industrial revolution, and the origins of economic growth in Britain. *Southern Economics Journal*, *74*, 50–70.

Morris, I. (2010). *Why the West rules*. New York: Ferrar, Strauss and Giroux.

North, D. C. (1990). *Institutions, institutional change and economic performance*. Cambridge: Cambridge University Press.

North, D. C., & Weingast, B. R. (1989). Constitutions and commitment: The evolution of institutions governing public choice in seventeenth century England. *Journal of Economic History*, *49*, 803–832.

North, D. C., Wallis, B., & Weingast, B. R. (2009). *Violence and social orders: A conceptual framework for interpreting recorded human history*. Cambridge: Cambridge University Press.

Nowak, M. (2011). *Supercooperators*. New York: Free Press.

Ormerod, P. (2001). *Butterfly economics*. New York: Basic.

Ostrom, E. (1990). *Governing the commons*. Cambridge: Cambridge University Press.
Pareto, V. (1935). *The mind and society [Trattato di sociologia generale]*. Brace, New York: Harcourt.
Parfit, D. (2011). *On what matters*. Oxford: Oxford University Press.
Penrose, R. (2011). *Cycles of time*. London: The Bodney Head.
Plott, C. R. (1967). A notion of equilibrium and its possibility under majority rule. *American Economic Review*, *57*, 787–806.
Popper, K. (1959). Prediction and prophecy in the social sciences. In P. Gardiner (Ed.), *Theories of history*. New York: Free Press.
Penn, E. (2009). A model of far-sighted voting. *American Journal of Political Science*, *53*, 36–54.
Pugh, C. C. (2002). *Real mathematical analysis*. Berlin: Springer.
Rader, T. (1972). *Theory of general economic equilibrium*. New York: Academic Press.
Richards, D. (1993). A chaotic model of power concentration in the international system. *International Studies Quarterly*, *37*, 55–72.
Richardson, L. F. (1922). *Weather prediction by numerical process*. Cambridge: Cambridge University Press.
Saari, D. (1985a). Price dynamics, social choice, voting methods, probability and chaos. In D. Aliprantis, O. Burkenshaw, & N. J. Rothman (Eds.), *Lecture notes in economics and mathematical systems*, No. 244. Berlin: Springer.
Saari, D. (1985b). A chaotic exploration of aggregation paradoxes. *SIAM Review*, *37*, 37–52.
Saari, D. (1995). Mathematical complexity of simple economics. *Notes American Mathematical Society*, *42*, 222–230.
Saari, D. (1997). The generic existence of a core for *q*-rules. *Economic Theory*, *9*, 219–260.
Saari, D. (2001a). *Decisions and elections: Explaining the unexpected*. Cambridge: Cambridge University Press.
Saari, D. (2001b). *Chaotic elections*. Providence, RI: American Mathematical Society.
Saari, D. (2008). *Disposing dictators, demystifying voting paradoxes*. Cambridge: Cambridge University Press.
Schofield, N. (1972a). Is majority rule special?. In R. G. Niemi & H. F. Weisberg (Eds.), *Probability models of collective decision-making*. Columbus, OH: Charles E. Merrill Publishing Co.
Schofield, N. (1972b). Ethical decision rules for uncertain voters. *British Journal of Political Science*, *2*, 193–207.
Schofield, N. (1975). A game theoretic analysis of Olson's game of collective action. *Journal of Conflict Resolution*, *19*, 441–461.
Schofield, N. (1977). The logic of catastrophe. *Human Ecology*, *5*, 261–271.
Schofield, N. (1978). Instability of simple dynamic games. *Review of Economic Studies*, *45*, 575–594.
Schofield, N. (1979). Rationality or chaos in social choice. *Greek Economic Review*, *1*, 61–76.
Schofield, N. (1980a). Generic properties of simple Bergson-Samuelson welfare functions. *Journal of Mathematical Economics*, *7*, 175–192.
Schofield, N. (1980b). Catastrophe theory and dynamic games. *Quality Quantity*, *14*, 529–545.
Schofield, N. (1983a). Equilibria in simple dynamic games. In P. Pattanaik & M. Salles (Eds.), *Social choice and welfare* (pp. 269–284). Amsterdam: North Holland.
Schofield, N. (1983b). Generic instability of majority rule. *Review of Economic Studies*, *50*, 695–705.
Schofield, N. (1984a). Existence of equilibrium on a manifold. *Mathematics of Operations Research*, *9*, 545–557.
Schofield, N. (1984b). Social equilibrium and cycles on compact sets. *Journal of Economic Theory*, *33*, 59–71.
Schofield, N. (1984c). Classification theorem for smooth social choice on a manifold. *Social Choice and Welfare*, *1*, 187–210.
Schofield, N. (1985). Anarchy, altruism and cooperation. *Social Choice and Welfare*, *2*, 207–219.

Schofield, N., & Tovey, C. (1992). Probability and convergence for supra-majority rule with Euclidean preferences. *Mathematical and Computer Modelling*, *16*, 41–58.
Schofield, N. (1999a). The heart and the uncovered set. *Journal of Economics Supplement*, *8*, 79–113.
Schofield, N. (1999b). A smooth social choice method of preference aggregation. In M. Wooders (Ed.), *Topics in mathematical economics and game theory: Essays in honor of R. Aumann*. Providence, RI: Fields Institute, American Mathematical Society.
Schofield, N. (1999c). The C^1-topology on the space of smooth preferences. *Social Choice and Welfare*, *16*, 445–470.
Schofield, N. (2002). Evolution of the constitution. *British Journal of Political Sciences*, *32*, 1–20.
Schofield, N. (2003). *Mathematical methods in economics and social choice*. Berlin: Springer.
Schofield, N. (2006). *Architects of political change*. Cambridge: Cambridge University Press.
Schofield, N. (2007). The mean voter theorem: Necessary and sufficient conditions for convergent equilibrium. *Review of Economic Studies*, *74*, 965–980.
Schofield, N. (2010). Social orders. *Social Choice and Welfare*, *34*, 503–536.
Schofield, N. (2011). Is the political economy stable or chaotic? *Czech Economic Review*, *5*, 76–93.
Schofield, N., & Schnidman E. (2011). Gridlock or leadership in U.S. Electoral Politics. This volume.
Schweitzer, F., et al. (2009). Economic networks: The new challenges. *Science*, *325*, 422–425.
Shafer, W., & Sonnenschein, H. (1975). Equilibrium in abstract economies without ordered preferences. *Journal of Mathematical Economics*, *2*, 245–248.
Shiller, R. (2003). *The new financial order*. Princeton: Princeton University Press.
Shiller, R. (2005). *Irrational exuberance*. Princeton: Princeton University Press.
Smale, S. (1966). Structurally stable systems are not dense. *American Journal of Mathematics*, *88*, 491–496.
Smale, S. (1974a). Global analysis and economics IIA: Extension of a theorem of Debreu. *Journal of Mathematical Economics*, *1*, 1–14.
Smale, S. (1974b). Global analysis and economics IV. Finiteness and stability of equilibria with general consumption sets and production. *Journal of Mathematical Economics*, *1*, 119–127.
Smith, A. (1984 [1759]). *The theory of moral sentiments*. Indianapolis, IN: Liberty Fund.
Strnad, J. (1985). The structure of continuous-valued neutral monotonic social functions. *Social Choice and Welfare*, *2*, 181–195.
Strogatz, S. (2004). *Sync*. New York: Hyperion.
Stern, N. (2007). *The economics of climate change*. Cambridge: Cambridge University Press.
Stern, N. (2009). *The global deal*. New York: Public Affairs.
Sunstein, C. R. (2006). *Infotopia*. Oxford: Oxford University Press.
Sunstein, C. R. (2011). *Going to extremes*. Oxford: Oxford University Press.
Surowiecki, J. (2005). *The wisdom of crowds*. New York: Anchor.
Taleb, N. N. (2007). *The black swan*. New York: Random.
Taleb, N. N., & Blyth, M. (2011). The black swan of Cairo. *Foreign Affairs*, *90*(3), 33–39.
Tataru, M. (1999). Growth rates in multidimensional spatial voting. *Mathematical Social Sciences*, *37*, 253–263.
Taylor, M. (1976). *Anarchy and cooperation*. London: Wiley.
Taylor, M. (1982). *Community, anarchy and liberty*. Cambridge: Cambridge University Press.
Thom, R. (1975). *Structural stability and morphogenesis*. Reading, MA: Benjamin.
Trivers, R. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, *46*, 35–56.
Trivers, R. (1985). *Social evolution*. Menlo Park: Cummings.
Walker, M. (1977). On the existence of maximal elements. *Journal of Economic Theory*, *16*, 470–474.
Watts, D. (2002). A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences*, *99*, 5766–5771.

Watts, D. (2003). *Six degrees*. New York: Norton.
Weitzman, M. (2009). Additive damages, fat-tailed climate dynamics, and uncertain discounting. *Economics*, *3*, 1–22.
Wilson, E. O. (2012). *The social conquest of Earth*. New York: W.W. Norton.
Zeeman, E. C. (1977). *Catastrophe theory: Selected papers, 1972–77*. New York: Addison Wesley.
Zipf, G. K. (1965). *Human behavior and the principle of least effort*. New York: Hafner.