The European Heritage in Economics and the Social Sciences
Series Editor: Jürgen Georg Backhaus

Alain Marciano
Giovanni Battista Ramello *Editors*

Law and Economics in Europe and the U.S.

The Legacy of Juergen Backhaus

Springer

The European Heritage in Economics and the Social Sciences

Volume 18

Series editor

Jürgen Georg Backhaus
Krupp Chair in Public Finance and Fiscal Sociology
University of Erfurt
Erfurt, Germany

The European heritage in economics and the social sciences is largely locked in languages other than English. Witness such classics as Storch's Cours d'Economie Politique, Wicksell's Finanztheoretische Untersuchungen and Geld, Zins und Güterpreise or Pareto's Trattato di Sociologia Generale. Since about 1937, partly caused by the forced exodus of many scholars from the German language countries and the international reactions to this event, English has become the undisputed primary language of economics and the social sciences. For about one generation, this language shift did not result in a loss of access to the European non-English sources. However, after foreign language requirements were dropped as entry pre-requisites for receiving the PhD at major research universities, the European heritage in economics and the social sciences has become largely inaccessible to the vast majority of practicing scholars.

In this series, we hope to publish works that address this problem in a threefold manner. An aspect of the European heritage in a language other than English should be critically documented and discussed, reconstructed and assessed from a modern scientific point of view, and tested with respect to its relevance for contemporary economic, social, or political discourse.

We welcome submissions that fit this bill in order to make the European heritage in economics and the social sciences available to the international research community of scholars in economics and the social sciences.

More information about this series at http://www.springer.com/series/5902

Alain Marciano · Giovanni Battista Ramello
Editors

Law and Economics in Europe and the U.S.

The Legacy of Juergen Backhaus

Springer

Editors
Alain Marciano
Université de Montpellier, LAMETA-CNRS
Montpellier
France

Giovanni Battista Ramello
Università del Piemonte Orientale
Alessandria
Italy

ISSN 1572-1744 ISSN 2197-5892 (electronic)
The European Heritage in Economics and the Social Sciences
ISBN 978-3-319-47469-4 ISBN 978-3-319-47471-7 (eBook)
DOI 10.1007/978-3-319-47471-7

Library of Congress Control Number: 2016954519

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Introduction

In July 2015, Jürgen Backhaus retired from his position as Professor at the University of Erfurt, after a long and rich career, started in 1970 at the University of Constance as an undergraduate student and that brought him all around the world. He was Professor of Public Finance at the University of Maastricht from 1986 to 2001 and, from 2001 to 2015, he held the Krupp Foundation Chair in Public Finance and Fiscal Sociology at the University of Erfurt.[1] Jürgen Backhaus published (and edited) tens of books and articles. He was, and still is, a scholar and a man of great immense culture. One of those rare scholars who is knowledgeable in so different fields that it is impossible to name all of them. But even within academia, Jürgen Backhaus is not only a scholar. He also played the role of a cultural entrepreneur. He launched, in 1994, the *European Journal of Law and Economics*, edited an important reference book, *Elgar Companion of Law and Economics*, and then had the idea of the *Encyclopedia of Law and Economics* that it is still an ongoing project involving a large number of contributors around the world, including ourselves as editors; he organized for decades one of the first and long-lasting European workshop in law and economics (first at the University of Maastricht and then at the University of Erfurt) and an interdisciplinary workshop in Heilbrönn. On the whole, he seemed to develop his scholarly activity as it were guided by the motto written by the Brazilian poet Vinicius De Moraes that said "Life is the art of encounter". We had the chance to participate in many of those encounters organized by Jürgen, we developed there a significant part of our scholarship and lastly we had the honor, in our turn, of being part of it as editors of journal and the encyclopedia.

Therefore, in addition to continuing the work and carry out the activities he created, we felt the need to pay him a tribute and to summarize what he did so far. Of course, this is a burden that can hardly be done by two sole individuals. Consequently, we had the chance to involve a number of well-known scholars that

[1]For a broader overview see his CV at https://www.uni-erfurt.de/fileadmin/user-docs/Finanzwissenschaft/Mitarbeiter/BackhausCVengllong-1.pdf.

in different ways have been connected to him. They enthusiastically accepted to contribute to this celebration. Hence, this book is thus the result of one other new encounter. The aim was not really to make an exhaustive presentation of the Backhaus's work, his contribution to law and economics or to characterize the field as it is nowadays, thanks to his contribution. We were rather interested in trying to evidence some crucial features of the law and economics movement. We believe that these features are also important to understand Jürgen Backhaus' work and conception of law and economics. The first chapter of this volume (co-written with Jean-Michel Josselin) provides further evidence of that. In this introduction, we summarize the three most important points.

The first important feature that characterizes law and economics, and that was too crucial for Jürgen Backhaus, relates to the European roots of law and economics. It is not simply to distinguish the European Journal of Law and Economics from a possible American counterpart nor to indicate the geographical location of its editors (him and Frank Stephen), that this name was chosen. Backhaus really believed that law and economics has a European legacy (see also Ramello 2016). More precisely, he believed that law and economics has a "continental" rather than Anglo-Saxon origins—let us note here that he also believed that there exists a particularly important continental tradition in public finance that leads to the development of public choice (see Backhaus and Wagner 2005). Of course, Jürgen Backhaus did not ignore that law and economics and public choice also developed in the USA—and in the rest of the world as well!—in a sort of dialogue or with continuous and repeated interactions. This is precisely what this volume evidences.

A second feature that fundamentally characterizes law and economics, and that also characterized Jürgen Backhaus' view of the economy, is precisely the crucial role of institutions and the importance of the interconnection between law, institutions, and economy. All the chapters gathered in this volume illustrate, in one way or the other, that no economy functions in an institutional vacuum and that institutions matter, to put it differently. Up to the point that institutions do not have the same meaning or role in different places. There is a form of relativism in law and economics that implies that a legal rule, an institution that exists in one country or in one environment emerged for specific reasons and that, as a consequence, one cannot copy it, transplant it easily in another environment. It could be said that law and economics is "comparative" in essence. This is also what these chapters tell us: we need to compare institutions to understand them.

This naturally and logically leads us to the third feature that characterizes this volume: the respective role of the market and the state. This is a vexed question, in particular in economics and in law and economics. If markets are efficient, then why should we need legal rules? And they are not, what kind of institutions do we need? The papers that are gathered in this volume all relate, in different ways, to these questions. They provide evidence that regulatory interventions of the state are far from costless but, on the other hand, that they are not useless. In particular, one aspect that is particularly important is the role of constitutions to ground and frame the economy.

All these aspects do not, however, cover all Jorgen Backhaus' research. Again Jürgen Backhaus is interested and knowledgeable in so many subjects and areas that it would be impossible to present all what he did.

Alain Marciano
Giovanni Battista Ramello

References

Backhaus, Juergen, and Richard E. Wagner. 2005. From Continental public finance to public choice: mapping continuity. *History of Political Economy*, 37 (Annual Supplement): 314–32.

Ramello, Giovanni B. 2016. The past, present and future of comparative law an economic, In *Comparative Law and Economics*, Eisenberg T. and Ramello GB. eds, Cheltenham, UK—Northampton, MA, USA: Edward Elgar Publishing.

The Law, The Economy, The Polity Jürgen Backhaus, A Thinker Outside the Box

Jean-Michel Josselin, Alain Marciano and Giovanni Battista Ramello

1 Introduction

It might surprise much people, including economists, if one states that economics has progressively but with a striking certainty turned into a formal science, axiomatized, that has not much to do with the political economy of its origins, that of David Hume or Adam Smith, and the other Scottish thinkers of the end of the eighteenth century. This is what James Buchanan already noted in 1958, when he and G. Warren Nutter decided to launch the Thomas Jefferson Center for the Study of Political Economy. Buchanan and Nutter were convinced that their discipline was drifting "away from its classical foundations as a component element in a comprehensive moral philosophy" (Buchanan 2007, 95) because "technique was replacing substance" (Buchanan 2007, 95) and because of the "increasing specialization of knowledge and scholarship" (Buchanan 1958, 5). Later, Buchanan and others as a sort of reaction "invented" Public Choice analysis, intended to be a modern form of political economy, with the objective to go back to the roots of the discipline. However, also Public Choice progressively became a form of rational choice politics and lost most of its political economy content. The same process took

The original version of this chapter was revised: For detailed information please see erratum. The erratum to this chapter is available at DOI 10.1007/978-3-319-47471-7_13

J.-M. Josselin (✉)
CREM-CNRS, University of Rennes, Rennes, France
e-mail: jean-michel.josselin@univ-rennes1.fr

A. Marciano
LAMETA-CNRS, University of Montpellier, Montpellier, France
e-mail: alain.marciano@umontpellier.fr

G.B. Ramello
Università del Piemonte Orientale, Alessandria, Italy
e-mail: giovanni.ramello@uniupo.it

A. Marciano and G.B. Ramello (eds.), *Law and Economics in Europe and the U.S.*,
The European Heritage in Economics and the Social Sciences,
DOI 10.1007/978-3-319-47471-7_1

place for the discipline focusing on the intersection of economics and the law officially founded in the USA in the Sixties. The law and economics to which Ronald Coase attached his name or the economic analysis of law that Guido Calabresi contributed to develop (see Marciano and Ramello 2014) turned again into a form of rational choice analysis of the law and legal institutions when Richard Posner transformed it into an economic analysis of law. Today, despite certain counter-tendencies, Posner's type of economic analysis of law remains dominant.

Where does Jürgen Backhaus stand in this picture? Of course, he is a public choice and a law and economics scholar. Moreover and undoubtedly, he played an important role in the law and economics and especially in its development in Europe. However, he played a crucial role in reconnecting the Chicagoan tradition of law and economics to its European roots and in turn the rational choice approach to political economy. Those roots, it will be further argued, were then somewhat lost. Jürgen Backhaus is one of those scholars putting back in evidence the legacy of the discipline to the important debates that took place in Europe during the eighteenth and the nineteenth century and that not only provided the ground for growing law and economics as a new discipline but also have been fundamental in shaping the physiognomy of modern Western societies.

At the beginnings, the efforts to make law and economics an autonomous discipline required to avoid any distraction and consequently brought the scholars to emphasize the methodological debate over any other issue (Ramello 2016). That with the consequence of neglecting sometimes quite evident features like the fact that two of the founding fathers, notably Guido Calabresi and Ronald Coase, are European.

Jürgen Backhaus has the merit of having brought again in evidence the link between Europe and the USA, His scientific and academic achievements are per se remarkable—the number of his articles published, and books authored or edited is impressive. His entrepreneurial activities are also worth being emphasized and, in a sense, are even more important. Likewise Henri Manne, who marketed law and economics in the USA at a time when no one really knew what law and economics was about and existed, Backhaus contributed to promote law and economics in Europe by organizing workshops, by editing the *European Journal of Law and Economics*, by editing a number of books, by educating a number of PhD students and lecturing around many countries, shortly by setting up the fertile humus on which to plant again the seeds of the discipline.

After a double higher education in law and in economics at the University of Konstanz and thus being at the same time exposed to both legal and economic scholarships, after a visit as post-doctoral Fellow at the Center for Study of Public Choice of the Virginia Polytechnic Institute and State University (1977–1978), Backhaus moved stably to the USA for the years 1980–1986 as an associate professor at the Department of Economics of Auburn University.[1] During this period,

[1]See https://www.uni-erfurt.de/fileadmin/user-docs/Finanzwissenschaft/Mitarbeiter/BackhausCVengllong-1.pdf.

he had the opportunity of matching together his double education and to grasp "law and economics" just when it began to become in the US an independent discipline. Once back to Europe, formerly at the University of Maastricht (1986–2001) in Netherlands and then at the University of Erfurt (since 2001), he started to diffuse the approach in Europe not only in his teaching activity, but also by undertaking the various activities mentioned above. A particularly important role has been played by the cycle of the Maastricht (then Erfurt) consecutive workshops in law and economics that for 27 years represented a stage in which the European scholars had the opportunity to present their work in progress and to refine it through the discussion with colleagues. An impressive number of the well-known members of the community definitively moved their steps, sometime their first steps, there.

The other noticeable accomplishment of Backhaus is the *European Journal of Law and Economics*, founded together with Frank Stephen (at the time rooted at the University of Strathclyde, Scotland) in 1994 and intended to become an important platform for hosting the scientific debate with a special "emphasis on European Community law and the comparative analysis of legal structures and legal problem solutions in member states of the European Community [… and] the new European market economies."[2]

If this was the focus of the journal, the methodological orientation once more reflected the broad cultural overview of Backhaus and the editors were alerting the reader of the need of a "pragmatic position informed by the history of law and economics as a scholarly discipline and the current challenges this discipline faces." (Backhaus and Stephen 1994).

Indeed, in methodological terms, this is the second point we would like to emphasize, Backhaus occupies a very specific place in the field. His work is not heterodox but not mainstream either. Likewise his cultural background, his form of law and economics actually evidences a way of going back to political economy, along the lines of what the great thinkers from the eighteenth and nineteenth century did. Along the lines of Buchanan, for instance, or Coase and Calabresi rather than along the lines of Posner or, for that matter, Gary Becker, he always tries to match the current development of the discipline to what the great thinkers already discussed. In the other important publishing adventure of Backhaus, the *Elgar Companion to Law and Economics*, once more a crossroads for favoring the blending between the American and the European scholarship of law and economics thanks to the contributions of members of the two communities, a significant part is devoted to introduce the work of important and somewhat neglected scholars like Cesare Beccaria, Friedrich List, Gustav von Schmoller, Rudolf von Jhering, among others.

We claim once more, and this is precisely how we want to pay tribute to Backhaus's work, that his scholarship adopted a novel political economy approach to law and economics and in this respect his contribution far exceeds the simple

[2]See the journal disclaimer at http://www.springer.com/economics/law+%26+economics/journal/10657.

diffusion of law and economics in Europe. In other terms, his commitment was not only aimed at fostering the consciousness of the European roots of the discipline but equally at nurturing a European approach to it. To this purpose, we analyze several of Backhaus's works and insist on a twofold aspect of his contributions to law and economics: the constant use of the history of economic thought and the reference to past but above all European thinkers. This double move did not allow only to rediscover economists, legal scholars, intellectuals who had been forgotten or neglected despite their importance, and who would probably have been ignored had it not been for his obstinacy. It also allowed him to give—or rather to reestablish a specific, European, identity to law and economics. Indeed, undoubtedly fascinated by the USA, Backhaus nevertheless definitely remains a continental European scholar fully aware of the importance of the legal and historical heritage of his origins. Indeed, at the core of this identity, one finds the need to embed economics in the larger pattern of interactions between the law, the economy, the polity, and the national conditions. Also of crucial importance is the close link between historical circumstances, the nature of a constitution and the very conception of the State. This has deeply influenced his approach to law and economics and contributed to its originality.

2 Political Economy and the Pure Science of Economics

The science of economics, or the scientific discipline of economics, that can be said to represent the core of our discipline can be viewed as "pure"—not only because it does not take into account institutions but essentially in the sense that it is primarily based on abstract models and because its purpose or objective is to propose universal and generic (see Wagner this volume on generic and specific explanations). The pure science of economics is based on the idea that one explanation can be given for all the behaviors that individuals adopt and decisions they make—independently from the circumstances and context in which those behaviors take place. Or, to put it in other words, the objective of a pure science of economics is to restrict its attention to behaviors that can be generalized or universalized.

From this perspective, one of the most typical instance of "pure" economics in the twentieth century is the one adopted by Paul Samuelson. The latter did argue that "[d]octrinal history shows that theoretical insight often comes from considering strong or extreme cases." (1955: 350) To him, one of these extreme polar cases was "[t]he grand Walrasian model of competitive general equilibrium" (1955: 350) while the other was his own "pure theory of government" (1954). Clearly, the consequence is that such a pure model as Samuelson's could give general and universal insights—in that case, the lesson was straightforward: when there exist public goods (or externalities), markets fail to allocate resources efficiently and governments must step in. However, the specifics were lost, as was noticed by some critics, in particular, Stefen Enke (1955), Julius Margolis (1955) and Tiebout (1956) who emphasized that, in the USA in the 1950s, most public goods were not pure but

local, in contrast to what Samuelson had assumed. In other words, Samuelson's representation of the economy was too pure and abstract and… unrealistic. Such lack of realism could not have been a problem if it were not for important consequences regarding the conclusions Samuelson had drawn. This was at the core of Tiebout's analysis: acknowledging the existence of local public goods could allow the economist to identify the institutional solutions that individuals find to organize the provision of these goods without the need for the intervention of the national state. More generally, one can say that the broad, pure and abstract, approach adopted by Samuelson hid the crucial ability of individuals to self-govern and to create institutions to organize such a self-governance—on this, see also, obviously, Buchanan's (1964) article on clubs.

The idea of self-governing societies—or, in other words, that individuals are able to devise mechanisms to govern their interactions—leads to another implication: the problems that individuals have and the solutions to remove those problems must come from the individuals themselves, and should not be imposed from the outside to the individuals in a top-down, organic, approach. In particular, one should not presuppose that collective welfare functions exist and public policies should not be based on them (Buchanan 1959).

Another instance, of particular interest from a law and economics perspective, is Gary Becker's analysis of crime and criminal behaviors (1968). Notwithstanding the rich and multifold perspective introduced by Beccaria (1764, 1994) on the same topic two centuries earlier (and recognized by Becker himself, 1968: 176, 209), providing the basis for a normative model including deterrence but also individual and social safeguards against the arbitrariness and the excess of punishment, Becker was then exclusively interested in proposing a model that could be used to understand the conditions under which an individual would decide to commit a crime. His approach of the problem was abstract, general and certainly not dependent on the specific situation of each criminal. And this allowed Becker to derive universal recommendations in terms of public policy—in a nutshell, it is better to fine criminals than to jail them.[3]

Both Samuelson's and Becker's models lead to "free-floating abstractions" (Boettke and Coyne 2005: 152), one of them being the abstract representation of the individuals who populate them. Indeed, this way of envisaging economic problems is possible because of the behavioral assumption that is used by those pure economists. It is not only that individuals are self-interested or even rational but that they are utility maximizers and that their utility functions are given.

Now, if individuals can really be fully characterized by a given utility function, then their choices are basically predetermined. These are not choices but mechanical decisions that any computer could make. This is precisely what

[3]It was precisely Becker's claim that economists would only need "thin" or "parsimonious" theoretical constructs, by contrast with sociologists, for instance, who need "thick" theories (on the distinction between "thin/parsimonious" and "thick" see Boettke and Coyne 2005; Boettke et al. 2006). And the advantages of an economic approach to crime is that it does not require a complex and detailed theoretical apparatus to explain phenomena.

Buchanan stressed in one of his most important articles, "What Should Economists Do?" (1964). To Buchanan, these mechanical decisions are not an object of study for economists but for engineers or computer scientists. The typical example of those decisions that are not choices—and of no interest for economists—are those of Robinson Crusoe: being alone on his island, Robinson Crusoe can select his most preferred objects in his environment; these choices involve no creativity because they are made by one isolated individual. Later, he would say that these behaviors are similar to the behaviors "rats" (Buchanan 1982) adopt in laboratory experiments. He argued that neoclassical economists were perfectly equipped to analyze these behaviors. But when Friday arrives on the island, then interactions between him and Crusoe generates creativity. Their choices are no longer the mechanical reactions of rats.

These two features are important to understand the differences between the pure science of economics and political economy. In effect, pure economics does not only assume that individuals always act in the same way but, much more important, that their behavior is independent from the situation or the context they face. The environment—including institutions—in which individuals act does not matter. This is precisely why universal models can be proposed. But, the situation changes if one adopts a more realistic approach of human behavior—not to say human nature—that accepts adaptation to the context, possibly failures, and certainly cognitive limitations. Let us note here that we are not necessarily talking of adopting a "behavioral approach" *à la* Sustein, Kahneman, Thaler or Jolls. One needs only to assume that individual preferences are not necessarily given and that they change, for instance, in interactions with other individuals (see Buchanan 1964). This implies that the context in which the interactions take place matters. Depending on whom they interact with or in which context the interaction takes place, individuals will not adopt the same behavior. And that, to understand and predict how individuals behave, it is necessary to "contextualize[e] the human condition" (Boettke and Coyne 2005: 152).

Similarly, another consequence is that institutions will not have the same meaning or purpose, and therefore cannot be evaluated in the same way, in different contexts. Similarly, as Richard Wagner (this volume) emphasizes, it may be interesting to speak of Common Law or Statute Law systems—as if these general categories, legal families, indeed existed—because it helps to gain certain insights on systems of law. But this kind of approach ignores that all systems of law are different. It ignores that, though systems of administrative law are slowly converging to a certain extent, they still remain "organically" different. Civil law countries are definitely not common law countries and are likely to remain so even if many attempts have been made at demonstrating the superiority of the latter over the former (La Porta et al. 2008). Legal origins have historical origins. As a consequence, students from civil law countries should have the opportunity to learn the economics and law that would help them understand also the place where they live. Furthermore, current research in the determinants of what a good life is beyond material growth will probably provide nice insights into the impact of legal systems

on quality of life, happiness being a primordial ingredient of it (Frey and Stutzer 2010) as well as the sense of belonging to a community.

This implies that the efficiency of these systems cannot be evaluated in a uniform way. Another example can be given by using an apparently simple market oriented institution, that is to say the trademark law. Trademarks emerged in very distinct cultural and legal settings—such as in ancient Greece, in the Roman Empire, and in ancient China for mentioning a few—and can thus be seen as a quasi-universal institution (Ramello and Silva 2006). Indeed its pervasiveness despite the many differences that characterized the societies and the economies in which it emerged can easily lead the observer to assert that its role is generally speaking "to promote economic efficiency" (Landes and Posner 1987: 265). Actually this has been the view of the standard law and economics literature which in accordance to the literature on asymmetric information regards marks and brands as signal for solving an adverse selection problem (Akerlof 1970; Landes and Posner 1987). However, despite appearances and a possible similar origin, the exploitation of trademark developed very differently in different contexts. For example, although both the Western world and China adopted trademarks, they did so with very different meanings, which can in turn account for their different ways of developing and enforcing them. In China, despite the spontaneous emergence and the venerable age of this trade device, the life of trademark had a totally different development, diverging from the endogenous market dynamics that characterized the Western world and that led toward the metamorphosis of the trademark in brand (Ramello 2006; 2016). To the point that despite almost a century China ago was pressured to amend Western-like laws because of the lobbying of Britain and notwithstanding the fact that China was a major exporter and trading nation, trademark-brand did not develop and did not play the same role as it did in many Western economies. These discrepancies persist in the present-days and to a great extent can be explained by the specific local reception and understanding of trademark (Alford 1995; Grinvald 2008).

Then, one clearly understands that what we need to appreciate is that the world is not homogenous. Indeed, behaviors and their meanings are not identical when one changes the conditions, culture, context in which individuals are. Therefore, the interactions that take place between individuals and, as a consequence, the institutions they devise to organize their interactions and solve the problems they face remain differentiated from one place to another. The world is characterized by a huge (institutional) variety and this must be recognized by economics. This thus means that there is not one answer to the question of what are "good"—or "bad"—institutions. One cannot give a definite answer, based on a pure, theoretical and abstract model. The answer depends on the context. What is or was good in China in terms of trademark was, or could not have been good in Western countries, and reciprocally. More broadly, as Buchanan (1959) and Buchanan and Stubblebine (1962) argued, an externality evidences a market failure—and has therefore to be removed—if and only if the individuals affected by the externality think that it has to be removed. This cannot be decided, from the outside, by omniscient economists (see Marciano 2013).

This also means, as a corollary, that one cannot easily transplant institutions from one context or culture to the other. This is a vexed question in law: can legal rules be transplanted from one country to the other? The answer could be positive if all the individuals were the same and if all institutions could mechanically cause the same effects in all contexts. This is not the case. For instance, a large part of the literature on the European institutional structure considers that it is a failure because it does not correspond to the American federal structure or to any other known form of federalism. This view means judging the European institutional structure without taking into account the historical context and the process that took the European nations from a dramatic conflict to a peaceful and relatively integrated structure. The European nations developed their own form of federalism, because they were trying to face specific problems (Josselin and Marciano 2007, 2004b). Indeed, federating nations cannot be the same as federating a nation, as it happened in the USA. Certainly, one can learn from the institutional evolution that led the USA from a confederation to a federation. But, this cannot be done by trying to impose the American model in Europe but rather by recognizing that the situations were and the institutions are different.

This is precisely what a political economy approach kicks in. Or, to put it in other words, to recognize the institutional variety and heterogeneity would lead economics back to political economy. As Simeon Djankov et al. noted: "economic analysis can move further by recognizing that different institutions are appropriate in different circumstances" (2003: 619). And, that the context matters, because human action is necessarily contextualized, is an insight that comes from political economists such as Mises, Hayek, Buchanan or Ostrom. Modern political economists (Boettke and Coyne 2005; Boettke et al. 2006) have also emphasized this point and concluded that one needs a theory—what they call "thin" or "parsimonious" theory—and "dirty" empirics. One needs "to combine the logical structure of economic reasoning with the rich institutional details of history and anthropological and sociological analysis" (Boettke and Coyne 2005: 157) to understand "unique institutional arrangements that structure the rules of the game and their enforcement in any particular historical setting." (Boettke et al. 2006: F309). Or, to put it differently, a pure—abstract—theory is not sufficient to understand what lies behind the general explanations and there always lies something behind. One always needs to refer to real-world circumstances in a sort of interplay between theories and empirics.

The trademark example is also useful because it tells us how important a comparative approach is if one really wants to gain insights in how institutions work. Similarly, if one for instance wants to improve our understanding of federalism, one should compare the American and European institutional evolutions (Josselin and Marciano 2004a). Hence it is not only the goal of comparative economics to deal with institutional diversity (Djankov et al. 2003: 619). Institutional studies in general carefully require a comparative approach and in a sense this is the only way for having a proper understanding—whether we look at similarities, or at differences (Ramello 2016). It is true for law and economics, and it is true for political economy. Let us note here that this comparison is twofold: it does not only

mean comparing different theoretical institutions—such the state or the market—but comparing different effective actual institutions. We are back to the idea of using real-world examples.

All the research agenda of Jürgen Backhaus extensively adopted this approach. This will be further argued in the next section.

3 Continental European Economic Thought Restored

One may start with a fundamental point of departure to understand Backhaus' work and conception of economics: he naturally emphasizes differences between the USA (or more generally the Anglo-Saxon tradition including England and its Commonwealth) and the European continent, in the same vein as in Frey and Eichenberger (1992).[4] To be clear, it is not simply that he claims that economics has European roots—which geographically and historically would not be particularly surprising and innovative—but also, that the (continental) European orientation in economics fundamentally differs from an American—or, more broadly, Anglo-Saxon—orientation. From this perspective, the most salient feature of each orientations is that the latter tends to be more "pure" than "political" while the former is more "political" than "pure." This in particular means that the European tradition is less abstract but more based on empirical findings that take into account and reflect cultures, history, national economic and legal conditions and heritage.

This is a leitmotiv throughout Backhaus's thinking and it is that the conceptualization of the relationship between the polity and the economy is crucial for defining the goal of the State and the scope of its activities, including public finance and the links between the government structure and fiscal policy. From this perspective, a first approach of what can be done privately, by the individuals, or what should be done publicly, by government consists in using an abstract theory without any or much connection with the real world and without empirical findings. This corresponds to the standard approach in public finance, as Backhaus emphasizes. Indeed, according to him, fiscal theorizing is largely independent from historical and institutional conditions. In the wake of Edgeworth and Pigou, taxation is conceived regardless of those circumstances and public finance is narrowly (to his view) instrumental for "appropriate intervention in the economic order" (Backhaus and Wagner 2005: 317, see also Backhaus and Wagner 2004) as if fiscal sociology narrowed down to an interference into a preexisting arrangement of the allocation of resources. By contrast with this "sacrificial" (ibid.) view of tax as an intervention in the marketplace, Backhaus is a proponent of an organic or interactive vision of individuals as citizens, states as participants within society, with markets organizing

[4]Let us note here that the importance of these European scholars had for Backhaus is another evidence that law and economics is a European field. Besides Coase or Calabresi, or Bentham and Beccaria (see Ramello 2016), there were also all the German scholars to whom Backhaus devoted a lot of work (see Sect. 4 below).

economic and social relations, and taxation a proactive instrument for promoting allocative efficiency and Pareto improvements in society. We are here quite far from government intervention against market failures or for distributive purposes based on (at best elusive) social welfare functions; Backhaus is adamant such that functions simply cannot be derived scientifically (Backhaus 1997) and he grounds his case on the path breaking work of Christian Wolff (1679–1754), as modern a scholar as he would have remained unknown and neglected without the constant obstinacy (shall we say doggedness?) of Backhaus. Wolff indeed questions the ability of discussions of equity and distributive justice to fit into a legal setting in which they will eventually have to be implemented. In other words, equity concerns should be built in their legal context and get an operational content from their inception.

Referring to the history of economic thought, and to history at all, allows Backhaus to teach us that some of our most famous writers indeed rooted their reflections in real-world circumstances. For instance, he quotes Wicksell complaining that our theorization of taxation dates too far back to times when the absolute powers of the seventeenth and eighteenth centuries monarchies in Europe would use Mercantilist reasoning to seek rent on a large scale and develop a competition "amongst the few" to try and rule the world through their empires. Mercantilism was indeed engrained in its geopolitical context, concurrently theorizing policies and feeding them with ideas and concepts, in an intricate pattern that is not so easily disentangled. Standard taxation theory has been mainly grounded on this historical and methodological setting. That way, what we could label "absolute taxation" became the norm in public economics teaching and research. In contrast, Backhaus has repeatedly explored a largely forgotten and ignored part of our European history, namely the tragic condition of German speaking people after the Thirty Years War (1618–1638). Forgotten history: the concepts of nation and empire are predominant in our (even scholarly) reasoning and it produces a kind of framing effect that prevents from thinking outside the box of national "big players," to use the phrase of Backhaus. Ignored history: the post Thirty Years War period was booming with ideas, real life experiments that should be as many objects of scientific curiosity, but have been largely neglected by teachers and scholars.

The pure tradition in public finance is typically represented by the standard Musgrave–Samuelson's approach of government finance and public goods based, as mentioned above, on the double assumption that individuals are self-interested utility maximizers but also that social planners are perfectly informed and more efficient than markets. Samuelson concluded that the intervention of the state was necessary. One knows that, in answer to Samuelson and Musgrave, Tiebout demonstrated how preference revelation is likely to be more effective through competition between small-scale jurisdictions. Although Tiebout's paper was rather abstract, his analysis had empirical foundations (see Singleton 2015).

The Tiebout setting can be illuminated by the "political order of small scale, competitive absolutisms" (Backhaus and Wagner 1987: 16) of the post war German speaking territories as they had been recognized by the Peace of Westphalia. The quasi-feudal structure of the more than 300 sovereign states offered an

unprecedented opportunity of government experiments in a context of high institutional fragmentation. Those fascinating (but in many respects appalling for the populations) historical circumstances have bred, under the pressure of events and the struggle for survival, an original conception of public finance through the Cameralist movement. The Cameralists were public administrators educated in law, economics, and management, serving local rulers who had to do with a competitive labor market with high mobility, and would provide local public goods to their citizens financed through the businesslike activities of necessarily enlightened rulers. Where government benevolence in large-scale absolutist states would depend on the goodwill of the ruler, in a Cameralist setting benevolence was a necessary ingredient of the sustainability or even survival of the regime. Ethical stances need not be called here: as Backhaus points out (Backhaus and Wagner 1987:17) "Cameralist thought was more solidly in favor of free trade or opposed to rent seeking than was mercantilist thought, but not because the Cameralists were better economists nor because they had higher morals, but because their princes faced different circumstances than did mercantilist kings."

Cameralist thinking provides insights into the nature of government that help go beyond traditional visions of representative governments (Blankart 1994). Club government is an appropriate instrument for understanding the wave of secessions and breakup of nations as they have been analyzed through spatial clubs theory (Alesina and Spolaore 2003; Josselin and Marciano 1999). The historical roots of economic thinking should never be neglected, Backhaus teaches us. The Cameralist experience also illuminates the necessity for good government to ground its action not only on pure economic incentive mechanisms, but also on an adequate and appropriate legal and institutional setting. Reading Backhaus makes you feel that economics is indeed law and economics.

4 Economics Is Law and Economics

This leads us to another important aspect of Backhaus' work that still relates to the double distinction between pure and political, on the one hand, and European and Anglo-Saxon on the other. Because cultures and context matter, as stressed above, institutions and systems of law still largely differ from one place to the other, shaped as they have been by language (e.g., emancipation from the Latin allows German lawyers like Otto von Gierke (1841–1921) to develop a specific Germanic school of law rooted into German history and legal traditions, see Backhaus 2005), customs often translated into law (e.g., a large part of the French civil code, see Josselin and Marciano 2002), imports of institutions (e.g., the Romans as well as Napoleon were familiar of the corresponding exports, being experts in imposing legal codes and systems of administration). This echoes the trademark example we gave in the first section of this paper.

This is why it is so important in Backhaus's view that national and linguistic traditions be taught, discussed, and put into perspective. For instance, a prominent

body of theory, "tightly tied to the German language area" (Backhaus, 2001: 459) is ORDO. Conceptualized in Freiburg by the economist Walter Eucken (1891–1950) and the lawyer Franz Böhm (1855–1977) at a time when law and economics could be taught in the same place to the same students, ORDO emphasized "the economic and legal order which provides the framework in which economic activity can take place" (ibid.). It has deeply influenced the making of the strikingly successful German economy after World War Two. With law and economics so intertwined, operational economic recommendations are both theoretically sound and most of all pragmatically relevant.

Subsidiarity is another instance of a concept that benefited from Backhaus' interpretation in terms of political economy. This is typically the kind of concept that escapes Anglo-Saxon standardization. A strange word long unrecognized by word processors, subsidiarity was reborn with the Maastricht Treaty on the European Union as an attempt at formalizing the allocation of prerogatives among levels of government. The historical circumstances dictated by the process of union would require guidance at a time when supranational entities, nations, and subnational governments would share a vast range of public actions and responsibilities. Oates' decentralization theorem (Oates 1999) did bring in a relevant economic insight into the criteria of allocation among levels of government, but would not articulate them into an institutional translation. Subsidiarity achieves it through the principle of delegation of tasks to the smallest functional unit of government (Backhaus 1999), then sequentially to the nearest functional unit (upwards or at the same level through cooperation) if the previous level cannot on its own efficiently fulfill the tasks. Of course, were tasks initially allocated to an intermediate body of government, then downwards subsidiarity would be triggered.

Subsidiarity takes on its first acceptation in Christian Wolff's (1679–1754) conception of a state welfare program whereby public intervention is only subsidiary to family and community bonds. That way, the efficiency of the state or larger community is optimized since it can direct its efforts and budgets toward the only actions that simply cannot be carried efficiently by the smaller community. Devolution to the smallest unit is not simply about taking advantage of its competence, it is also about lifting the burden that would inappropriately lay on the larger unit: "In order to optimize the performance of the larger political entity, primary liability for the solution of problems lies with the smallest functional unit" (Backhaus 1997: 138). Backhaus has substantially contributed to put Christian Wolff back into light. He has shown that Wolff's view of institutions is pragmatic, theoretically sound, and fully flexible since it allows reallocation of prerogatives in any direction (upwards, downwards, lateral) whenever new circumstances dictate it. Economics and constitutional law are thus intertwined: "pure" economic principles must be rooted in appropriate bodies of government as well as public law must allow the flexibility of institutions in order for them to be responsive to new challenges in public intervention.

Backhaus has always been adamant that (law and) economics scholars try to reach back to the founding fathers or thinkers who first illuminated, admittedly in a different historical setting, the very policy questions they currently try to tackle. In

the field of federalism for instance, we were quite surprised to find out that such a brilliant mind as Hans Kelsen (1881–1973) had been totally neglected by the economic literature on decentralization and more generally on the organization of prerogatives in a multilevel governmental setting. Kelsen is of course renown and celebrated by constitutional and public law scholars on the European continent, but had strangely remained absent from public finance and even public choice debates. A constitutional lawyer by profession, a constitutional writer by circumstances, Kelsen was an outstanding thinker who could have fitted Backhaus's view of Cameralists (mutatis mutandis), men making the state within the state.

5 Conclusion

Economics has largely evolved from political economy to a pure science. In this process, some general insights were gained but others were lost. The need to go back to a political economy approach—empirical, historical and comparative—is crucial. Jürgen Backhaus is one of those scholars who repeatedly and consistently carried out the message that the making and implementation of a constitution is definitely constitutive of the nature of the state. Conceiving economic intervention, allocative, or redistributive, outside these considerations can go against the existing state construction, or be simply non-implementable, or even illegal, or, to become legal, induce a distortion of the legal construction that presides to the very conception of the state as it has been chosen by citizens in a democracy rooted in its history. Backhaus has also been a key craftsman for the emancipation of a European way of thinking (law and) economics. The Maastricht (then Erfurt) workshop in law and economics has been a place for talking, learning, and enjoying academic debates. Young scholars met there Backhaus's friends from the USA and Europe, and they learned to emancipate their thinking from mainstream.

References

Akerlof, G.A. 1970. Uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84 (3): 488–500.

Alesina, A., and E. Spolaore. 2003. *The Size of Nations*. Cambridge: Cambridge University Press.

Alford, W.P. 1995. *To Steal a Book is an Elegant Offense. Intellectual Property Law in Chinese Civilization*. Stanford: Stanford University Press.

Backhaus, J.G. 1997. Christian Wolff on subsidiarity, the division of labor, and social welfare. *European Journal of Law and Economics* 4: 129–146.

Backhaus, J.G. 1999. Subsidiarity. In *The Elgar Companion to Law and Economics,*, edited by Backhaus JG, 136–143. Cheltenham, UK: Edward Elgar.

Backhaus, J.G. 2001. Introduction: The intellectual future of the concept of Staatswissenschaften. *European Journal of Law and Economics* 4: 129–146.

Backhaus, J.G. 2005. Otto von Gierke (1841–1921). In *The Elgar Companion to Law and Economics*, 2nd ed., 519–521. Cheltenham, UK: Edward Elgar.

Backhaus, J.G., and F.H. Stephen. 1994. The Purpose of the European Journal of Law and Economics and the Intentions of Its Editors. *European Journal of Law and Economics* 1: 5–7.

Backhaus, J.G., and R.E. Wagner. 1987. The cameralists: A public choice perspective. *Public Choice* 53: 3–20.

Backhaus, J.G., and R.E. Wagner, ed. 2004. *The Handbook of Public Finance*. Boston, MA: Kluwer Academic Publishers.

Backhaus, J.G., and R.E. Wagner. 2005. From continental public finance to public choice: Mapping continuity. *History of Political Economy* 37: 314–332.

Beccaria, C. 1764, 1994. *On crimes an Punishments and other Writings*. Cambridge: Cambridge University Press.

Blankart, C.B. 1994. Club governments versus representative governments. *Constitutional Political Economy* 5: 273–285.

Becker, G. S. 1968. Crime and punishment: an economic approach. *Journal of Political Economy* 76(2): 169–217.

Boettke, Peter J., and Christopher J. Coyne. 2005. Methodological individualism, spontaneous order and the research program of the Workshop in Political Theory and Policy Analysis. *Journal of Economic Behavior & Organization* 57: 145–158.

Boettke, Peter J., Christopher J. Coyne, John Davis, Francesco Guala, Alain Marciano, Jochen Runde, and Margaret Schabas. 2006. Where Economics and Philosophy Meet: Review of the Elgar Companion to Economics and Philosophy With Responses from the Authors. *Economic Journal* 116 (June): F306–F325.

Buchanan, J.M. 1958. The Thomas Jefferson Center for Studies in Political Economy. *The University of Virginia News Letter* 35(2): 5–8.

Buchanan, J.M. 1959. Positive economics, welfare economics, and political economy. *Journal of Law and Economics* 2 (October): 124–38.

Buchanan, J.M. 1964. What Should Economists Do? *Southern Economic Journal* 30: 213–222.

Buchanan, J.M. 2007. *Economics from the Outside in: "Better Than Plowing" and Beyond*. Texas A&M University Press.

Buchanan, J.M., and W.C. Stubblebine. 1962. Externality. *Economica* 29(116): 371–84.

Djankov, Simeon., Edward Glaeser., Rafael La Porta., Florencio Lopez-de-Silanes., and Andrei Shleifer. 2003. The new comparative economics. *Journal of Comparative Economics* 31: 595–619.

Enke, S. 1955. More on the misuse of mathematics in economics: a rejoinder. *Review of Economics and Statistics* 37(2): 131–133.

Frey, B., and R. Eichenberger. 1992. Economics and economists: A European perspective. *American Economic Review* 82: 216–220.

Frey, B., and A. Stutzer. 2010. *Happiness and economics. How the economy and institutions affect human wellbeing*. Princeton: Princeton University Press.

Grinvald, L.C. 2008. Making Much Ado about Theory: The Chinese Trademark Law. *Michigan Telecommunications and Technology Law Review* 15: 53–106.

Josselin, J.M., and A. Marciano. 1997. The paradox of Leviathan: How to develop and contain the future European state. *European Journal of Law and Economics* 4: 5–21.

Josselin, J.M., and A. Marciano. 1999. Unitary states and peripheral regions: A model of heterogeneous spatial clubs. *International Review of Law and Economics* 19: 501–511.

Josselin, J.M., and A. Marciano. 2002. The making of the French civil code: An economic interpretation. *European Journal of Law and Economics* 14: 193–203.

Josselin, J.M., and A. Marciano. 2004a. Federalism and Conflicts over Principalship. Some Insights into the American Constitutional History. *Constitutional Political Economy* 15: 281–304.

Josselin, J.M., and A. Marciano. 2004b. Federalism and subsidiarity, in national and international contexts. In *The Handbook of Public Finance*, edited by Backhaus JG and Wagner RE, 477–520. Dordrecht: Kluwer Academic Publishers.

Josselin J.-M., and A. Marciano. 2007. How the court made a federation of the EU. *Review of International Organisation* 2: 59–75.

La Porta, R., F. Lopez-de-Silanes., and A. Shleifer. 2008. The economic consequences of legal origins. *Journal of Economic Literature* 46: 285–322.

Landes, W.M., and R.A. Posner. 1987. Trademark law: An economic perspective. *Journal of Law and Economics* 30: 265–309.

Marciano, A. 2013. Why Market Failures Are Not a Problem: James Buchanan on Market Imperfections, Voluntary Cooperation, and Externalities. *History of Political Economy* 45(2): 223–254.

Marciano, A., and G.B. Ramello. 2014. Consent, Choice and Guido Calabresi's Heterodox Economic Analysis of Law. *Law and Contemporary Problems* 77: 97–116.

Margolis, J. 1955. A comment on the pure theory of public expenditure. *Review of Economics and Statistics* 37(4): 347–349.

Oates, W. 1999. An essay on fiscal federalism. *Journal of Economic Literature* 37: 1120–1149.

Ramello, G.B. 2006. What's in a sign? Trademark Law and Economic Theory. *Journal of Economic Surveys* 20: 547–565.

Ramello, G.B. 2016. The past, present and future of comparative law an economics. In *Comparative Law and Economics,* edited by Eisenberg, T. and G. Ramello. Cheltenham, UK —Northampton, MA, USA: Edward Elgar Publishing.

Ramello, G.B., and F. Silva. 2006. Appropriating Signs and Meaning: The Elusive Economics of Trademark. *Industrial and Corporate Change* 15: 937–963.

Samuelson, Paul A. 1954. The Pure Theory of Public Expenditure. *Review of Economics and Statistics* 36: 387–389.

Samuelson, Paul A. 1955. Diagrammatic Exposition of a Theory of Public Expenditure. *Review of Economics and Statistics* 37: 350–357.

Singleton, J. 2015. Sorting Charles Tiebout. *History of Political Economy* 47(annual suppl. 1): 199–226.

Tiebout, C. 1956. A Pure Theory of Local Public Expenditures. *Journal of Political Economy* 64: 416–424.

Economic Efficiency and the Law: Distinguishing Form from Substance

Richard E. Wagner

Scholars of law and economics have long been fascinated with and intrigued by claims that legal processes promote economic efficiency. Richard Posner's initial (1973) edition of *Economic Analysis of Law* is a treatise on jurisprudence wherein Posner claims that the entire body of common law rulings can be rendered coherent by recognizing that those rulings promote economic efficiency. The subsequent literature spewed theoretical seeds in several directions. One notable direction concerned whether statute law was also economically efficient (Wittman 1989, 1995; Backhaus 1998), something that Posner originally denied. Another concerned was whether efficiency resulted from the intention of judges or was a systemic product of the common law process, a topic that was central to James Buchanan's (1974) long review of Posner (1973). Other efforts replaced Posner's equilibrium framework with a framework that entailed evolutionary development (Rubin 1977, 1982; Priest 1977). Yet another line of thought claimed that the search for legal efficiency was misplaced because the emphasis was better placed on the stability of the legal framework, for it is legal invariance and not the adaptability of law to changing circumstances that facilitates economic efficiency (Epstein 1980, 1995; Rizzio 1980).

This essay explains why claims regarding the economic efficiency of legal arrangements are problematical in any case. In short, those claims mostly confound the form of an argument with its underlying substance. Economic efficiency is a feature of a particular economic model, the model of competitive equilibrium. In evolutionary and other nonequilibrium models, efficiency is undefined. Efficiency pertains to the form and not the substance of an economic model. Efficiency claims are instances of Paretian derivations whereby a logical-sounding argument is set forth to justify what cannot truly be demonstrated but is desired by the speaker all

R.E. Wagner (✉)
Department of Economics, George Mason University, 3G4, Fairfax, VA 22030, USA
e-mail: rwagner@gmu.edu
URL: http://mason.gmu.edu/∼rwagner

A. Marciano and G.B. Ramello (eds.), *Law and Economics in Europe and the U.S.*,
The European Heritage in Economics and the Social Sciences,
DOI 10.1007/978-3-319-47471-7_2

the same, as Pareto (1935) sets forth and as Patrick and Wagner (2015) and Wagner (2016) elaborate. Furthermore, the efficiency claim presumes the existence of something that should surely be the task of analysis to establish, namely whether there exists some universal point of agreement common to all members of a society. In this respect, Vilfredo Pareto once asked how it could be sensible to speak of maximizing happiness for a community when happiness for the wolves required eating the lambs while happiness for the lambs required the avoidance of being eaten. Even more, to speak of common law or statute law is to speak of some social whole without regard to how that whole is constituted and without regard to the relationships and interactions among the constituent elements of that whole. In other words, there are numerous possible versions of a common law process, with differing operating properties, that can reside within the general rubric of common law.

1 A Quick Review of Some Efficiency Claims

Posner's (1973) *Economic Analysis of Law* was fundamentally a treatise on the coherence of common law. Posner's self-adopted burden was to show that the body of common law rulings can be rendered coherently by recognizing that common law operates to render judgments in favor of economic efficiency. Someone who understood the principles of economic efficiency and who subsequently read through the body of common law rulings would see economic efficiency as the common thread that unites those myriad rulings across time and place. A jurisprudential theory of common law can thus be constructed around economic efficiency as an imperative of the common law process. Subordinate to this primary claim was the claim that statute law did not have the same efficiency properties. Hence, there would seem to be some tendency for societies to be wealthier the more fully legal relationships are governed by common law relative to statute law.

Before distinguishing between form and substance with respect to claims that common law promotes economic efficiency, it may be helpful to set forth briefly a few illustrations that Posner (1973) uses to illustrate the efficiency claim. For instance, in his chapter on property law, Posner explained that a railroad owed a duty of care to people, trespassers, who crossed the tracks at recognized crossing points, but owned no such duty to people who crossed those tracks elsewhere. In contrast, railroads owed a duty of care to trespassing cattle at all times. The difference in treatment between people and cattle reflects economic efficiency, Posner asserts, by invoking claims about the comparative costs of preventing accidents. It would be very costly for a railroad to prevent people from crossing tracks everywhere. Furthermore, trains would mostly encounter people at recognized crossing points, where a duty of care would entail relatively low cost. In contrast, it would be very costly for ranchers to fence their land to keep their cattle from wandering across railroad tracks. Collisions between trains and cattle can be prevented more cheaply by placing the duty of care on railroads rather than placing it on ranchers.

Children were treated differently than adults in this application of economic analysis to law, and this difference was likewise described as illustrating economic efficiency. Children have not yet acquired the full range of cognitive sensibilities that adults mostly have, though children have more acute sensibilities about danger than do cattle or sheep. To this threefold distinction between adults, children, and cattle, the legal doctrine of attractive nuisance reflects recognition of the three levels of cost. Railroads did not owe a duty of care to children who crossed or played on railroad tracks. In this, children were treated like cattle. But such things as railroad turntables would understandably be attractive to children, and here railroads owed a duty of care to watch for children. Hence, legal principles regarding trespass over a railroad's property reflected a threefold distinction among adults, children, and cattle that reflected comparative efficiencies in the avoidance of accidents.

Perhaps nowhere is the claim in support of economic efficiency as providing coherence to common law rulings more fully in evidence than in Posner's treatment of torts, a treatment that Landes and Posner (1987) amplify and extend. The instrument for doing this is Judge Learned Hand's formulation in *United States v. Carroll Towing Co*. [159 F.2d 169 (2d Cir. 1947)]. This case concerned an unattended barge that had broken loose from her moorings in New York Harbor. The legal issue was whether the owner of the barge was negligent in leaving the barge unattended, and thus liable for damages. In his ruling, Hand asserted that negligence in this case depended on three considerations: (1) the likelihood that the barge would break loose from her moorings, (2) the likely damage that would result if the barge did break loose, and (3) the burden involved in ensuring that the barge did not break loose. Hand summarized his judgment by invoking a piece of algebra that has become a staple formulation in the law and economics literature. In particular, liability for negligence was said to result if $B < PL$, where B is the burden or cost of preventing the barge from breaking loose, L is the damage that would result from breaking loose, and P is the likelihood or probability that the barge would break loose.

Judge Hand's formulation of liability in *Carroll Towing* can be readily apprehended according to Coase's (1960) later presentation of efficiency and liability. The owner of the barge would be liable for damages under either of two circumstances. Under one circumstance, the owner of the barge had the right to let his barge wander in the harbor unless the other users of the harbor could buy the barge owner's agreement to tether his barge. Under the alternative circumstance, the owner of the barge had no such right unless he could buy the agreement of the other users of the harbor to leave his barge untethered. Regardless of the initial circumstance, the Hand formula leads to the same assignment of liability as expressed in Coase's subsequent formulation. If the expected cost from the damage caused by a wandering barge exceeded the cost of keeping the barge tethered, the Hand formula yields that outcome regardless of the initial locus of property rights.

Just as this formulation provides a useful framework for thinking about this and related situations, it is also apparent that the three variables in what has come to be called the Hand Formula reflect observer judgments and not unambiguously objective magnitudes. For instance, Hand noted in his decision that it would be

unreasonable to expect a barge to be attended continuously, even in a crowed harbor during war. But Hand also noted that the barge had been left unattended for 21 h, which he declared to be an excessive length of time. While the variables in Hand's formula are judgments and not facts, those judgments can serve as a useful heuristic for organizing thought about this and similar situations. In this respect, Landes and Posner (1987, pp. 96–107) examine 14 cases with respect to the Hand Formula.

The first two of those cases can be used to illustrate how the Hand Formula might be used to lend coherence through economic efficiency to contrasting judgments about liability in tort cases, and with Wagner (1992) providing further amplification. In *Hendricks v. Peabody Coal Co*. [115[1].App. 2d 35, 253 N.E.2d 56 (1969)], a 16-year-old boy dove into an abandoned strip mine that had filled with water. He was injured upon hitting a submerged shelf. The court ruled for the plaintiff, noting that the abandoned mine could have been enclosed by a fence at a cost of between $12,000 and $14,000. This cost was low relative to the potential damage from diving into the water. In their review of this case, Landes and Posner (p. 97) declared that "the court was on safe ground in concluding that the defendant had failed to use due care." In *Adams v. Bullock* [227 N.Y. 208, 125 N.E. 93 (1919)], by contrast, the court ruled against the plaintiff. In this case a 12-year-old boy was swinging an eight-foot long wire as he was walking across a bridge that passed over an electric trolley track. The boy's wire touched the trolley wire, burning the boy. In ruling against the plaintiff, the court held that the injury was an "extraordinary casualty, not fairly within the area of ordinary prevision."

The contrary rulings in the two cases can be reconciled within the framework of the Hand Formula, Landes and Posner argued. In *Hendricks v. Peabody Coal*, the likelihood that people would find the water-filled mine an attractive swimming hole was high, while the cost of fencing off the hole was relatively small. For *Adams v. Bullock*, the reverse relation held. It is not likely that people would be encountered who were dangling wires while crossing over a trolley track. Furthermore, it would be comparatively expensive to cover all bridges to prevent such situations. The claim on behalf of common law efficiency is that if a set of rulings is separated between those the plaintiff wins and those that the defendant wins, economic efficiency will be on the side of the winners whether plaintiffs or defendants.

It is easy enough to understand why many economists have been attracted to claims that legal processes supported economic efficiency. Yet it is also possible to find cases that seem to point clearly in the contrary direction. Consider two of the many cases that Huber (1988) examines. In one case, a man tried to mount a 16.5 in. tire on a 16 in. rim. To get the tire to hold to the rim, the man had to inflate the tire to 48 lb of pressure per square inch. After the tire expanded when the air heated up after driving for some time, the tire exploded, crashing the car and injuring the man. The man sued the manufacturer on the grounds that the company had not warned him against the dangers of overinflating the tire and of putting the tire on the wrong-sized rim. The man won his case. In a second case, a teenaged boy was burning a candle in his room. Wanting to add aroma to his room, the boy poured cologne over the burning candle, engulfing himself in flames in the process.

The manufacturer was ruled liable for the boy's burns, on the grounds of failing to warn about the flammability of cologne. Consideration of such cases as these almost unavoidable leaves one to wonder whether the claim of economic efficiency is a reasonable scientific finding or a metaphysical ordering principle that speaks particularly strongly to economists.

2 Common Law Efficiency: Science or Metaphysics?

To claim that common law rulings reflect economic efficiency requires a theorist to claim to be able to distinguish objectively what is efficient from what is inefficient. With respect to the preceding set of cases, there might be intuitive plausibility at work in making the aforementioned distinctions regarding the assignment of court verdicts based on economic calculation. Still, intuition is a subjective quality, and people can differ in their judgments. Those who lose most cases probably think they had the better case. The central claim of economic theory, moreover, is that efficient economic outcomes are not subject to determination by outside parties. Rather, efficient outcomes are conclusions drawn from an understanding of the operating properties of a particular institutional arrangement.

Within an institutional arrangement governed by the legal principles of private property, freedom of contract, and liability for harms and damages, the internal logic of the market economy is that transactions will continually move resources from employments that are less highly valued by resource owners to employments that are more highly valued. If such a market economy is conceptualized in equilibrium terms, no unexploited gains from trade will exist. It is not, however, possible to cite any set of economic observations as corresponding to a state of equilibrium within the context of economic theory. Indeed, it is almost surely the case that actual societies are always operating within a nonequilibrium environment because continual experimentation and change is a normal feature of modern life. To recognize this situation, however, is to point to some alternative theoretical framework that conceptualizes a process that operates over some duration of time rather than to conceptualize a state of affairs that exists as some particular instant of time, as Wagner (2010) explains.

Consider the manner in which economists derive cost functions from production functions. To start, production is conceived as a process by which inputs are combined to produce some output. If X is output and a and b are inputs, $X = f(a, b)$. The inputs a and b must be combined to produce X. Typically, those inputs can be combined in various ways, some of them more costly than others. In this respect, it is typically assumed that producers seek to minimize the cost of producing any particular output. To do this requires that they select a combination of inputs such that the ratio of the marginal products of the inputs equals the ratio of input prices. From these production relations, cost functions are readily derived. Perhaps the most notable feature of a cost function is that it creates a separation between situations that are possible and situations that are impossible. A cost function

describes a relationship between cost and output, and with that function derived from assuming that a firm minimizes the cost of producing output. For any given output, any cost measure above the cost function is possible while any measure below that function is impossible.

The cost function is an imaginary construction that is developed by facing a hypothesized firm with different production functions and input prices. Yet economic theory is based on the presumed congruence of those functions with observable reality. Yet there is no way that such a boundary can be observed. As a logical matter, it is impossible to assert that an actual cost of production is below the boundary. Any observed cost of production must either lie on the boundary or above it. Why, then, locate it on the boundary when that boundary is impossible to locate? To declare that common law is economically efficient is to locate legal processes as operating at the boundary between possibility and impossibility. This claim might be a reasonable metaphysical ordering principle, but it cannot be claimed to be a refutable statement about the world of actual experience.

What makes this boundary claim seem reasonable is that it corresponds with reasonable intuitions about human nature within the institutional arrangements governed by private property and freedom of contract. Those arrangements create positions of residual claimacy wherein some people own the residual between the revenue a firm derives from selling its output and the expenses it incurs in hiring the inputs necessary to produce that output. It is reasonable to think that people who receive that residual, which can be negative as well as positive, will prefer larger to smaller residuals. If a residual claimant can develop a lower cost method of producing the same output, that claimant will have strong motivation to shift to that lower cost method. Recognition of this motivation does not demonstrate that production in market economies always takes place along the boundary. Indeed, the simple observation that many firms fail and undergo reorganization is evidence that not all firms operate along the boundary. Still, residual claimacy is an institutional arrangement that yields a plausible basis for thinking that the institutional arrangements of a market economy have a strong tendency to induce firms through experimentation to gravitate toward least-cost input combinations. By extension, something similar could be said about legal processes if they were governed by the same institutional framework.

But legal processes are not governed by that type of framework. Neither are contemporary economic processes for that matter. Much economic activity is organized through governmental entities which operate through a budgetary process that operates in a significantly different fashion from residual claimacy. Buchanan (1969) explains that the cost of an action is the value of the highest-valued alternative action that the chooser rejects in choosing the preferred action. Cost and choice are reciprocals, as the essays in Buchanan and Thirlby (1973) elaborate. The cost of a choice typically differs when it is made under residual claimacy then when it is not. With respect to legal processes, for instance, residual claimacy might lead two commercial litigants to settle a dispute because they are residual claimants to their legal expenses. Should the plaintiff be a governmental agency, however, principles of residual claimacy are not in play. Should the public agency settle the

case, there is no residual for executives or owners of the agency to capture. Whatever expenses of litigation might be saved by settlement will be swept back into the agency's budget. Cost is different for a public litigant than it is for a private litigant, due to the absence of residual claimacy for public litigants. A public litigant who settles a case rather than going to trial has no residual to claim. Either that unclaimable residual is returned to the Treasury or is spent on other activities preferred by agency executives.

Popper (1959) locates the boundary between science and metaphysics according to whether a claim is falsifiable or just verifiable. While Popper's demarcation has received much criticism on various grounds since he first advanced it that demarcation point to a significant distinction is all the same even if falsifiability is incapable of being implemented. When the various controversies are cleared away, what perhaps remains is recognition that there are two forms of verification, one subjective and one objective, or at least intersubjective. In *Carroll Towing*, for instance, the categories in Judge Hand's formulation are subjective in that they pertain to Judge Hand's sense of the matter. No external and objective appraisals of B, P, and L were presented that would command universal assent by their objective quality. This does not mean that judgment is arbitrary in the sense that anything is possible. It does, however, mean that reasonable people can reach different judgments regarding the same situation.

This recognition has implications for claims about legal efficiency. Consider again Posner's illustration of railroads, people, and cattle and his argument that owing a greater duty of care to cattle than to people illustrates economic efficiency at work. With respect to Posner's claim, Tullock (1980) points out that Posner's claim is not accompanied by evidence that speaks to his claim. For instance, Posner asserts that it would be less costly for a pedestrian to choose a path that avoided crossing a railroad track than it would be for locomotive engineer to watch continually for passengers. This might be so, but no evidence is presented on the point. The efficiency claim is not a hypothesis that can be tested, but is rather a logical implication of a prior presumption that common law rulings reflect economic efficiency. If someone presumes that common law rulings reflect economic efficiency, it must be concluded that it is relatively more costly for railroads to exercise care toward passengers than toward cattle. Yet a locomotive engineer who is watching for cattle will unavoidably see pedestrians at the same time, so the marginal cost of watching for pedestrians is zero. Recognition of the joint cost character of watching for cattle and pedestrians refutes the claim that the differences in the duties of care reflect economic efficiency. It would seem to be the case that the desire to treat economic efficiency as giving coherence to the body of common law comes first, and with observations pertaining to particular rulings woven around that metaphysical ordering principle. It is here where Pareto's distinction between logical and nonlogical action becomes relevant to appraising the claims on behalf of common law's ability to promote economic efficiency.

3 Paretian Derivations and Efficiency Claims

It is easy enough to accept that these efficiency claims because they sound reasonably reasonable. But perhaps this reasonable quality reflects a preceding willingness to believe the claim. In this respect, Pareto (1935) advanced the vital distinction between logical and nonlogical action, and with Backhaus (1978) exploring some of the public choice implications of Pareto's distinction. It should be noted at the start that this distinction is not a distinction between rational and irrational, though a number of commentators on Pareto have asserted that it is. For Pareto, all action was rational. In this respect, Pareto would surely have agreed with Szasz's (1961) formulation that mental illness was largely a myth created to make it easier for the speakers to make their speech. For instance, an elderly and wealthy widow with four children who she believes are doing little more than waiting for her to die so they can inherit her wealth, may use her wealth to endow an orphanage, or worse, an asylum for unwanted dogs and cats. If the children can have her declared mentally incompetent, they can contest their mother's will and inherit her fortune.

There might be nothing wrong with the widow's mental faculties. Certainly, leaving her fortune to establish a foundation to support stray animals rather than supporting her adult children who have led shiftless lives is in no way evidence of mental incapacity. Indeed, it could well be evidence of acute mental capacity in recognizing shiftlessness in her children in conjunction with their anticipations of receiving hefty inheritances. In contesting the will, moreover, the adult children could not expect to find a sympathetic judge or jury to support their desires to live shiftless and profligate lives. To be successful in their pursuit of inheritances, the adult children would have to develop derivations that resonated with the sympathies of those would decide about their contestation of their mother's will.

This situation fits nicely Pareto's distinction between logical and nonlogical action. That difference has nothing to do with some actions being rational and others being irrational. The difference is rather due to different environments in which action takes place, with some environments eliciting action of the logical type and other environments eliciting nonlogical types of action. Basically, logical action is the domain of action within market settings while nonlogical action is the domain of action within political and religious settings. All action aims at improving an actor's situation relative to what that situation would otherwise have been. But there are different environments in which action occurs, and the substantive content of rational action plays out differently between those environments that elicit logical action and those that elicit nonlogical action. The former environments correspond to notions of inspection and experience goods, while the latter environments correspond to credence goods.

In market settings, people take actions to alleviate uneasiness they sense. They might be hungry and seek a place to eat. They might be unhappy with their old television and want to get a new one. Whatever the object at which the actor aims, the actor is engaged in a scientific-like process of forming and testing hypotheses. In some cases the qualities of goods can be reasonably well gauged by inspection, as

in looking over items at a salad bar. In other cases, those qualities require some period of experience with the good, as illustrated by a television set. In either case, buyers form images of what they are looking for, and can compare vendor offerings with the prices they are asked to pay.

Furthermore, vendors are in open competition with one another within this particular type of market setting. With experience goods in particular, vendors will have to overcome some understandable reluctance of buyers to buy a product when they cannot determine a product's qualities until after the purchase has been made. There are numerous things vendors do to overcome that reluctance. One important thing is the development of reputation. Products and producers that develop strong reputations for delivering reliable quality will face less resistance in selling experience goods. That reluctance can be lowered further by such practices as allowing returns within 30 days, and with this practice being less costly to producers of reliable products. In other words, logical action for Pareto corresponds to a scientific-like setting where vendors advance claims about the ability of their products to satisfy buyer desires. Potential buyers can test those claims by choosing to buy one product over another, and in a context where various practices and conventions have emerged through the efforts of vendors to overcome possible buyer reluctance, especially with relatively high-priced experience goods.

Not all arenas within which people act conform to the scientific-like setting of logical action where people perform experiments with their resources, choosing outcomes based on those experiments. With respect to the earlier illustration of a widow and her adult children, logical action would pertain to an environment where the children were exploring different options for caring for their mother, making a choice based on the evidence they accumulate. With nonlogical action, by contrast, the desired end is first chosen and the challenge for action is to get the required other people to support that desired end. The children want their mother's estate for themselves and not as a foundation for orphaned animals. But they need support from other people, who have their own values and constraints, to be able to achieve this end. To claim openly and forthrightly that they want their mother's fortune for themselves is unlikely to muster much support. To combine some psychiatric examination with a declaration of wanting to do good for their mother will surely be more effective in getting control of their mother's estate. Reason is still in play with nonlogical action, but it operates within a different environment from ordinary market environments.

Political and religious arenas, Pareto recognized, mostly involve environments where evidence cannot be acquired and acted upon. Voters, for instance, cannot choose their desired politicians or policies. In this alternative environment, political vendors likewise recognize that listeners will not subject a candidate's claims to scientific-like tests because the nature of the settings renders this impossible. Political competition thus revolves around the creation of ideological images by candidates and parties, seeking to construct images that resonate well with voter sentiments. Political candidates are in the same position as the adult children who wanted to gain control of their mother's estate, and had to construct an ideological image that would resonate positively with the sentiments of those who controlled that outcome.

With respect to religion, numerous efforts have been made to render religious belief a matter of logical action by explaining why a person must believe in the existence of God. Some of those arguments invoke a chain of causation that is traced back to an original uncaused cause, and with that uncaused cause pointing to God. Others have made use of probabilistic arguments, as illustrated by Pascal's wager in which a rational gambler would choose to believe in God based on calculations of expected value. In these types of arguments, logic-based arguments are invoked to convince the listener to embrace a belief in the existence of God. These formulations seek to reduce belief to logic, and with the employment of the relevant logic forcing belief upon an otherwise skeptical person.

In sharp contrast was Anselm's approach to God's existence, which is summarized by the title of Barth (1960), *Anselm: Fides Quaerens Intellectum*. In this instance, belief is the point of departure and not the destination, for it is faith seeking understanding. This is Pareto's approach to nonlogical action. Belief precedes action, it is not generated through action, as when repeated satisfaction with a particular product creates brand loyalty. With logical action, an action is taken based on a hypothesis about the consequences of that action. With nonlogical action, a belief or desire creates a corresponding action. To avoid appearing arbitrary, the taker of any particular action must give logical-sounding reasons even though the correct order runs from desired outcome to supporting reasoning.

4 Institutional Arrangements: Generics Versus Specifics

Scholarly controversies over common versus statute law, or over democratic versus authoritarian political regimes, typically operate at a highly aggregate level of discourse. Legal systems are thus distinguished according to whether they operate according to common or civil law systems. Similarly, political systems are distinguished according to whether they are democratic or authoritarian, with democratic meaning that some political officials are selected through election. While this manner of approach is readily susceptible to statistical analysis, the meaningfulness of such efforts is also questionable, especially if there are particular details inside those systems that do significant work in channeling outcomes in particular directions.

With respect to democratic polities, for instance, it is common to distinguish between presidential and parliamentary systems and to use statistics to reveal differences in average values between those classes of regime. This procedure is genuinely informative, however, only to the extent the generic difference in form accounts for the substantive differences between the regimes. Within each class of regime, however, enormous differences are possible in many respects, and those differences may be responsible for the observed differences among regimes. For instance, a democratic regime that operated under a constitutional requirement that all revenue must be raised by a flat-rate tax on all income without exemptions or exceptions would surely exhibit significantly different characteristics than one

where legislative majorities can do whatever they choose with respect to taxation, even if this leads to a majority of the population being exempt from tax. Whether a regime is presidential or parliamentary may pale in significance besides the systemic properties through which revenues are raised.

The same types of issues pertain to the economic analysis of legal systems. The standard distinction between common law and civil law is purely generic, and yet the most significant lines of analysis might require efforts to plumb institutional details regarding those systems. If the devil truly resides in the details, as a piece of ancient wisdom remarks, this will truly be the case. Hogue (1966) and Berman (1983) show that common law practices originated in an environment that more closely resembled what people mean when they speak of free and open competition wherein judges had to attract custom than is true these days. The same term "common law" is used to describe wide differences in the practical arrangements through which law is generated. At an earlier age, common law emerged through decisions of judges and juries as these were subsequently rendered coherent by such codifiers and systematizers as Blackstone (1979 [1765–1769]). Legislation resided in the far background of the common law process in Blackstone's time. These days, the requirements of legislation reside in the foreground.

In the early days, the common law process was polycentric. The formation of law was a bottom-up process, and with scholars like Blackstone seeking to find and explain the unity that existed among the rulings across different courts. These days, the common law process is monocentric, and with conflicting rulings across jurisdictions being something to be eliminated by a higher court or through legislation, as against being an indication of local differences in relevant sensibilities. To say this is not to say that the old ways were better, but is only to note that common law is a generic or formal term that does not prescribe some particular process. These days, for instance, jurors are silent during a trial and are given parameters to stay within in reaching their determination. At an earlier time, jurors could ask questions during a trial and were participants in the conduct of a trial. Again, to say this is not to make a judgment but is only to note that analysis at the generic level might as much obscure reality as it reveals it.

5 Some Concluding Remarks

There is a bidirectional relationship between law and economics. From one direction, producing law is an ordinary economic activity, as Hogue (1966) explains in his analysis of the entrepreneurial construction of particular causes of action. At any instant there is a social division of labor, wherein such occupations as judges and lawyers emerge through the same processes of economizing action as do all other occupations in society. Among other things, changes in the pattern of occupations and activities through time or among places should be amenable to the same

principles of economizing action. Humans are both cooperative and quarrelsome creatures, and law is necessary both to harness the gains from cooperation and to restrain the destructive power of quarrelsomeness unleased. At the same time, however, changes in particular details regarding legal arrangements can confer advantages on some while imposing disadvantages on others. The intersection of law and economics provides a fascinating and fecund vantage point for observing the human drama in all of its glory and malice, and which Jürgen Backhaus's writings and editorial activity have done much to illuminate.

References

Backhaus, J.G. 1978. Pareto and Public Choice. *Public Choice* 33: 5–17.
Backhaus, J.G. 1998. Efficient Statute Law, vol. II. In *The New Palgrave Dictionary of Economics and the Law,* edited by P. Newman, 24–28. London: Macmillan.
Barth, K. 1960. *Anselm: Fides Quaerens Intellectum*. Richmond, VA: John Knox Press.
Berman, H.J. 1983. *Law and Revolution: The Formation of the Western Legal Tradition*. Cambridge: Harvard University Press.
Blackstone, William. 1979 [1765–1769]. *Commentaries on the Laws of England*. Chicago: University of Chicago Press.
Buchanan, J.M. 1969. *Cost and Choice*. Chicago: Markham.
Buchanan, J.M. 1974. Good Economics. Bad Law. *Virginia Law Review* 60: 483–492.
Buchanan, J.M., and G.F. Thirlby, eds. 1973. *L.S.E. Essays on Cost*. London: Weidenfeld Nicolson.
Coase, R.H. 1960. The Problem of Social Cost. *Journal of Law and Economics* 3: 1–44.
Epstein, R.A. 1980. The Static Conception of the Common Law. *Journal of Legal Studies* 9: 253–275.
Epstein, R.A. 1995. *Simple Rules for a Complex World*. Cambridge, MA: Harvard University Press.
Hogue, A.R. 1966. *Origins of the Common Law*. Bloomington: University of Indiana Press.
Huber, P. 1988. *Liability: The Legal Revolution and Its Consequences*. New York: Basic Books.
Landes, W., and R.A. Posner. 1987. *The Economic Structure of Tort Law*. Cambridge, MA: Harvard University Press.
Pareto, V. 1935 [1923]. *The Mind and Society: A Treatise on General Sociology*. New York: Harcourt Brace.
Patrick, M., and R.E. Wagner. 2015. From Mixed Economy to Entangled Political Economy: A Paretian Social-theoretic Orientation. *Public Choice* 164: 103–116.
Popper, K. 1959. *The Logic of Scientific Discovery*. New York: Basic Books.
Posner, R.A. 1973. *The Economic Analysis of Law*. Boston: Little, Brown.
Priest, G. 1977. The Common Law Process and the Selection of Efficient Rules. *Journal of Legal Studies* 6: 65–82.
Rizzio, M.J. 1980. Law amid Flux: The Economics of Negligence and Strict Liability in Tort. *Journal of Legal Studies* 9: 291–318.
Rubin, P.H. 1977. Why is the Common Law Efficient? *Journal of Legal Studies* 6: 51–63.
Rubin, P.H. 1982. Common Law and Statute Law. *Journal of Legal Studies* 11: 205–223.
Szasz, T.S. 1961. *The Myth of Mental Illness: Foundations of a Theory of Personal Conduct*. New York: Harper & Row.
Tullock, G. 1980. Two Kinds of Legal Efficiency. *Hofstra Law Review* 9: 659–669.
Wagner, R.E. 1992. Crafting Social Rules: Common Law vs. Statute Law, Once Again. *Constitutional Political Economy* 3: 381–397.

Wagner, R.E. 2010. *Mind, Society, and Human Action: Time and Knowledge in a Theory of Social Economy*. London: Routledge.
Wagner, R.E. 2016. *Politics as a Peculiar Business: Insights from a Theory of Entangled Political Economy*. Cheltenham, UK: Edward Elgar.
Wittman, D. 1989. Why Democracies Produce Efficient Results. *Journal of Political Economy* 97: 1395–1424.
Wittman, D. 1995. *The Myth of Democratic Failure*. Chicago: University of Chicago Press.

Bootleggers and Baptists in the Garden of Good and Evil: Understanding America's Entangled Economy

Dima Yazji Shamoun and Bruce Yandle

1 Introduction

Since the early 1970s, the U.S. economy has experienced significant and rarely interrupted growth in federal regulation. We see this, for example, in Fig. 1, which reports the annual count of pages of new and modified rules published in the U.S. government's *Federal Register*, across the years 1940 through 2014. The *Federal Register* is the official daily chronicle for all newly proposed and final rules produced by the federal government. As readily observed, the 1970s set a high bar for later growth.

Part of this sudden expansion of rules is explained by the creation of new regulatory agencies. The U.S. Environmental Protection Agency was established in 1970 along with major environmental statutes that required development of rules affecting air, water, and land pollution. The U.S. Consumer Products Safety Commission and Occupational Safety and Health Administration were also established in the early 1970s as new federal statutes were passed that supplanted state and local regulatory dominance and increased the pace of activity for older regulatory agencies.

The authors are Associate Director of Research at the Center for Politics & Governance and Lecturer at the Economics Department, University of Texas at Austin and Dean Emeritus, College of Business and Behavioral Science, Clemson University and Distinguished Adjunct Professor of Economics, Mercatus Center at George Mason University, respectively.

D.Y. Shamoun (✉)
Economics Department, University of Texas at Austin, Austin, USA
e-mail: DShamoun@utexas.edu

B. Yandle
Mercatus Center, George Mason University, Fairfax, USA
e-mail: yandle@bellsouth.net

A. Marciano and G.B. Ramello (eds.), *Law and Economics in Europe and the U.S.*,
The European Heritage in Economics and the Social Sciences,
DOI 10.1007/978-3-319-47471-7_3

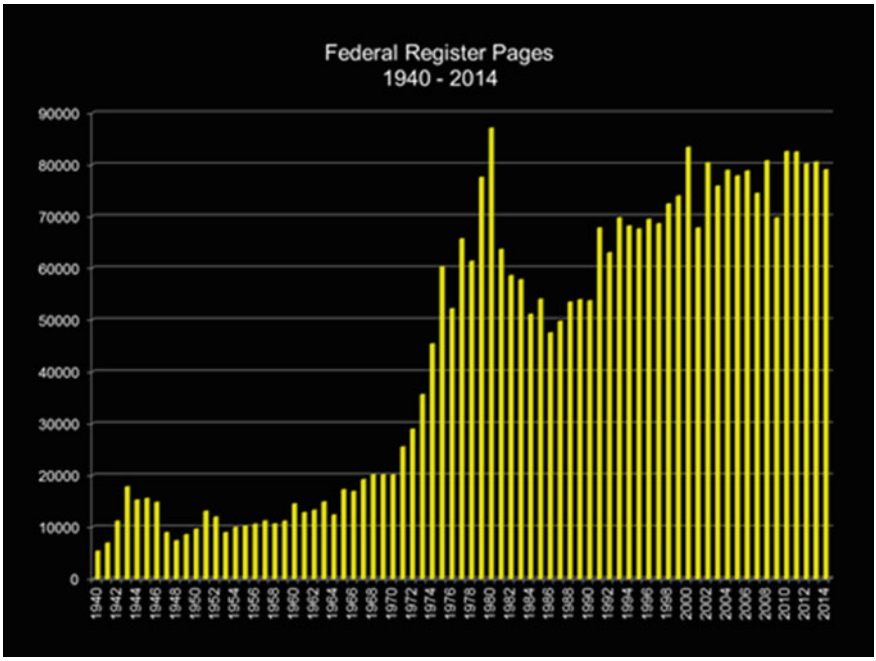


Fig. 1 Annual count of *federal register* pages, 1940–2014

Growth in the economy itself may have also stimulated growth in regulation. It is possible that a larger—and more complex—economy somehow requires additional federal rules. To illustrate this possibility, Fig. 2 reports the count of *Federal Register* pages divided by real GDP. In fact the 1970s regulatory surge is even more pronounced when adjusted for GDP.

Along with technical change that spurred economic growth and, perhaps, the demand for federal rules, the 1970s marked the final formation of a national market for consumer goods accommodated by network television (Yandle 2010). Participants in national markets called for federal rules to replace the multiplicity of state and local regulations that previously regulated locally produced goods and services. Thus, there may have been other exogenous stimuli that led to high growth in central government regulation.

Volumes have been written analyzing the rise of the U.S. administrative state, and countless journal articles, along with journals to publish them, have emerged since the 1970s partly in an effort to explain and predict the surge of U.S. regulatory activity. It is not our intention to review or add to this literature. Instead, we assume a simpler task. We seek to explain why, within the regulatory surge, technology-based command-and-control (or CAC) regulations (i.e., regulations that are implemented using CAC instruments) became the instrument of choice as the U. S. economy became entangled with federal rules (Smith et al. 2011). We wish to explain what made CAC instruments so attractive, relative to other regulatory

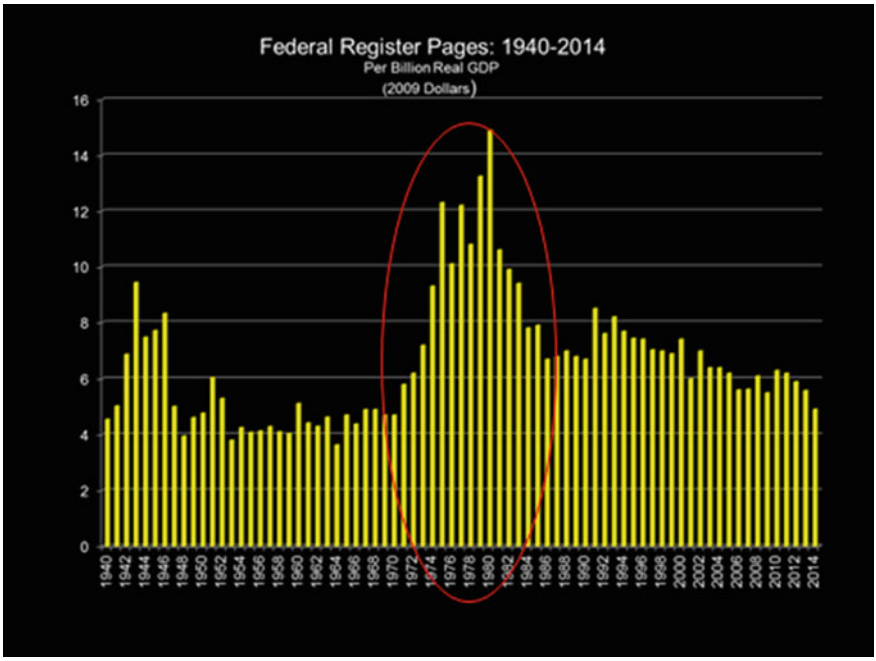


Fig. 2 Annual count of *federal register* pages divided by U.S. real GDP

instruments, e.g., taxation, the setting of performance standards, or use of property rights, that might have been chosen.

Our explanation applies Public Choice concepts to highlight the behavior of economic and political agents when operating in the political arena. As the title to our paper implies, we refer to the political arena as the "Garden of Good and Evil." We see the political arena as a commons where interacting participants, when regulating, can take actions that, at the margin, may improve the wealth of the nation, a good outcome, or as a place where political agents act in ways that reduce overall well-being, which we refer to as an evil outcome. From within the Public Choice toolkit, we select and enrich Yandle's (1983) Bootlegger/Baptist theory of regulation and use it as a foundation for explaining why CAC methods became the instrument of choice.

In applying Yandle's theory, we will show how CAC regulations can embody elements of good (welfare enhancing) and evil (welfare reducing) in the same regulation and by doing so become extraordinarily attractive to participants in the Garden of Good and Evil. We note that when good and evil are packaged in the same regulation, durable coalitions form to reduce the political cost of forming regulation. We also note that it is impossible, a priori, to draw firm conclusions regarding a bundled regulation's overall welfare effects. Put another way, we offer a positive analysis of the political economy, one that aims to explain the way the

world works, as opposed to a normative argument that claims to evaluate desirability of outcomes.

This chapter is organized as follows: In Sect. 2, we develop additional background on the rise of U.S. regulation and review theories of regulation that have been offered by historians, political scientists, and economists in an effort to explain regulatory behavior. We also discuss the choices made when politicians are selecting which regulatory instrument to apply as they construct a regulatory apparatus; we go on to describe some of the forces that favor one instrument over another. Section 2 also introduces the Bootlegger/Baptist theory.

Section 3[1] enriches the Bootlegger/Baptist theory and focuses on the politician's challenge: How to satisfy the regulatory demands of diverse interest groups who are essential to his political success? To enrich the story, we draw on the work of Bueno de Mesquita and Smith (2012), which provides a useful framework for a finer grain analysis of politician and interest group interaction. The political solution to the politician's challenge calls for development of a hybrid regulatory package, one that satisfies the regulatory demands of both public and private interests (Shamoun 2013). Drawing on a metric developed at George Mason University's Mercatus Center, Sect. 3 also provides evidence on the frequency of CAC regulation. The frequency and related effects, which we report, support our contention that the U.S. has become an entangled economy (Wagner 2009) because of the fertile soil the Bootleggers and Baptists have found in the Garden of Good and Evil. We conclude the paper with brief final thoughts.

2 Choosing How to Regulate

The explosion of U.S. federal regulation that began in the 1970s revealed a fundamental challenge to politicians who were constructing the legislative blueprints that instructed regulatory agencies as to how to build the detailed regulation that followed. The challenge had to do with which *instrument* to select for achieving the regulatory goal. Would the pending regulatory focus be better implemented by imposing price controls or higher taxes and fees, assigning property rights, regulating entry and exit, and setting performance standards; or would the rules describe changes in how, when, where, and by whom goods and services may be produced?

Regulations can be dichotomized into two broad *categories*: social or economic. Social regulations generally address issues relating to health, safety, security, and the environment. They are normally focused on a narrow issue, e.g., carbon emission, but their jurisdiction can extend to multiple industries, such as energy and transportation (e.g., coal plants, automobile manufactures, etc.). Economic regulations, on the other hand, deal with entry, quality of service, fares, prices, and rates of return in specific industries, such as transportation, communications, water,

[1]The ideas in this section originated in Shamoun (2013).

electricity, and natural gas (Brito and Dudley 2012). Both of these categories of regulations can be fulfilled using several instruments, such as, economic incentives, performance standards, property rights, or by implementing CAC approaches. While market-based and performance-based instruments dictate the outcome of the regulation, CAC instruments dictate the *means* by which the regulatory outcome is to be achieved. Historically, economic regulation has been attained by the former, and social regulation the latter.

When selecting regulatory instruments, the differences within the two categories are by no means trivial. For example, scholars were in broad agreement that setting performance standards, for example, to reduce carbon emissions by 30 % over some baseline, using any approach that might be chosen by regulated parties, was the low-cost regulatory approach. Performance standards did not require centralized authorities to ferret out dispersed knowledge and then decide once and for all, for example, how refrigerators should be built or electricity-generating plants operated. The use of performance standards provided competitive discovery incentives, made it possible to introduce new technologies when they were developed, and gave bottom-line incentives to minimize cost. Performance standards forced regulators to focus on outcomes—were the regulations really performing?—instead of on means —were the plants and machines being appropriately operated according to some standards?

Employing economic incentives was another theoretically attractive regulatory choice. For example, defining once and for all a limited number of emission allowances and allowing trade in them to constrain carbon emissions could also induce discovery of lower cost control techniques. Alternately, the use of emission fees and taxes would do the same thing. Setting a price on any unwanted activity would encourage economic agents to economize and reduce their harmful behavior. Again, regulators would not have to get into the engineering business; they could focus on outcomes, not means.

Finally, regulators could go the high-cost route of implementing CAC instruments, limiting discovery incentives. They could attempt to become experts in designing electricity generators, steel mills, food processing plants, automobiles, air conditioners, and all other major pollution sources and develop engineering standards that specify how emissions would be reduced. This, in fact, became the dominant U.S. regulatory instrument. We seek to explain why.

2.1 The Choice of Regulatory Instruments: A Political Choice

Whether an activity is regulated by means of CAC instruments or by means of economic- or performance-based instruments is a political choice. As mentioned, protection of air and water quality can be achieved by specifying the kinds of machinery to be operated by polluting firms—a CAC approach—or protected by

imposing discharge fees or taxes on all polluters and other users of the scarce environmental assets, which would be an economic-based approach. Furthermore, air and water quality can be protected by limiting the entry of firms and organizations whose activities will consume environmental quality. In other words, environmental regulation could be designed in ways that satisfy the definition of economic-based instruments. In addition, the environment can be protected by establishing property rights that empower right holders to bring a legal action against polluters who impose cost on them without their permission. Similar approaches—fees, taxes, fines, and rights—could be devised to manage safety, health, and consumer protection. When writing regulation-spawning laws, politicians make choices.

Despite the feasibility and efficacy of economic- and performance-based approaches in implementing social regulations, politicians have begun to consistently rely on the use of CAC instruments to achieve their goal. Therefore, the increase of social regulations has been accompanied by the increase of CAC instruments. A record reflecting the distribution of regulations by category is shown in Fig. 3, which reports the budgeted expenditures for U.S. regulatory agencies across the years 1960–2015. The chart shows data for three regulatory categories: economic, social, and transportation safety administration, which is the expenditure for air travel security that followed 9/11. Obviously, when it comes to the use of tax

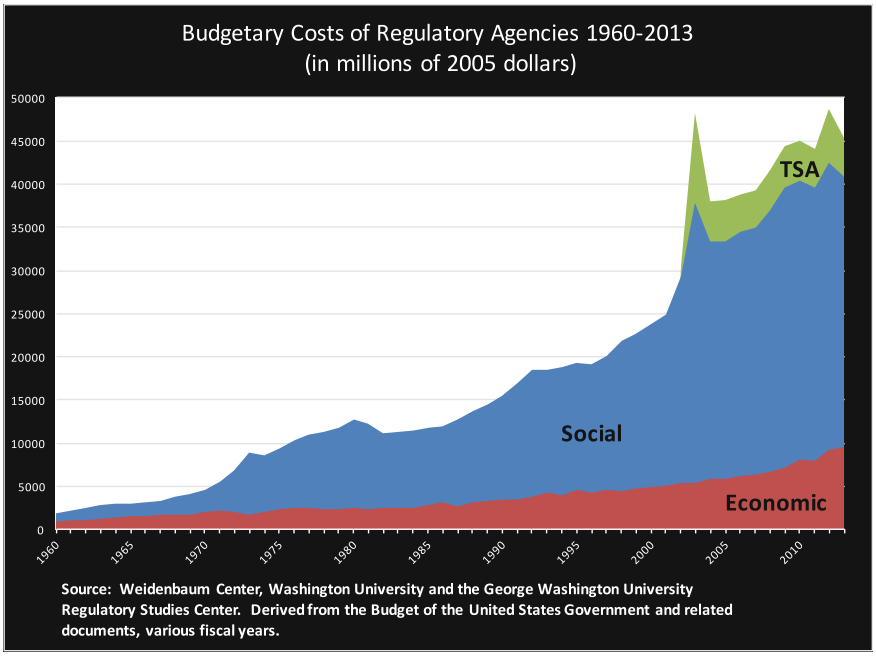


Fig. 3 Regulatory agency budgets for three regulation categories

dollars, politicians strongly favor social regulations, which gives us an idea of how often CAC instruments are invoked when crafting regulations.

Public Choice logic suggests why this would be the case. CAC instruments enable politicians to more accurately predict and target which firms, technologies, and industries will be most affected, and which constituencies will bear the greatest cost and which will receive the larger benefit. In addition to targeting, CAC approaches can achieve regulatory outcomes while at the same time hiding the incremental cost associated with the regulation. Consumers of air quality will not receive a monthly bill, nor will statements of charges be sent systematically to industrial firms that discharge waste into rivers, produce faulty consumer products, or operate unsafe workplaces. And of course, it is possible to have uniform rules and differential enforcement, which may make this approach all the more attractive to politicians.

2.2 What Theory Explains Regulator Behavior?

These last few statements raise questions regarding what motivates regulators. Can their actions best be explained by appealing to Public Choice logic, which suggests that politicians, like other normal people, are motivated by hope of personal gain, whether it be continued employment, higher future income, or social esteem? Or, alternatively, are political regulators driven primarily by an altruistic desire to serve the broad public interest to make the world a better place? There are, after all, competing theories to consider. Over the decades, political scientists, historians, and economists have struggled to develop a positive theory of regulation. Each theory has its strengths and its weaknesses (Brito and Dudley 2012).

The oldest of these explanations is called the Public Interest Theory and is associated erroneously with the name of Arthur Cecil Pigou (1920), a noteworthy English economist who, among other topics, focused on addressing the problem of social cost. Pigou described countless situations where, in his opinion, private action imposed largely uncompensated costs on the public at large. These included drivers of cars that wore out city streets, women who worked instead of caring for their children, and the more typical cases of factories that belched smoke on clothes drying at a nearby laundry. All of these could theoretically be addressed by an all-knowing political body. Pigou (1920) later indicated that no political body would behave in ways to serve the public interest in such matters, but would instead be swayed by special interest influence.

Under the Public Interest theory, government regulators are seen as working diligently to serve the broad public interest. Not motivated by the prospects of personal gain, these regulators work to correct market failures that lead to monopolized markets, environmental degradation, and shoddy consumer products, and also to protect the wealth, health, and safety of low-paid workers. Whether couched in terms of externalities, underprovision of public goods, or information

asymmetries, the work of the public-spirited regulator is seen as being on the side of angels, while recognizing that regulators are still human.

While close observers of political action will likely agree with Pigou that special interest influence does seem to prevail when government spending programs are debated or tax policy considered, there is still a deeply committed group, especially among environmentalists, who act as though they believe regulators will more generally serve the public interest, which they claim is also their interest. Generally speaking, supporters of environmental CAC instruments do not see the process generating those rules as being just another part of transaction politics where politicians deliver regulations, just as they might shuffle to their supporters increases in particular packages of defense spending, all in exchange for political support.

Dissatisfaction with the overall usefulness of Public Interest theory for explaining political behavior, led to the development of a second theory of regulation, the Capture Theory, which was elaborated by Bernstein (1977) and Kolko (1963). Capture Theory can be thought of as beginning with a committed Public Interest regulator who truly seeks to provide a cost-minimizing or welfare-maximizing regulatory bundle. But in the course of seeking information about the problem to be addressed, becomes acquainted with, let us say, industry officials who work diligently to provide useful data and analysis to the politician. Unwittingly, perhaps, the politician becomes captured or unduly influenced by industry. According to the theory, this is how the public interest is compromised.

If one must choose either Capture Theory or Public Interest Theory for explaining the behavior of politicians engaged in a long series of regulatory transactions, one might be tempted to name Capture Theory the superior model. After all, the theory recognizes the economic value of regulations that can be provided politically, while simultaneously accepting the reality of transactional politics. Yet while it may be more useful, it suffers from a major shortcoming. There are often competing interest groups that wish to influence political outcomes.

For example, when fuel producers, engine manufacturers, and transport companies are involved with rules intended to reduce nitrogen oxide emissions from heavy trucks, it is not clear what the regulatory outcome would be. Will the nitrogen oxide rule reflect primarily the wants and interests of the fuel producers more than those of the transport companies and the engine manufactures? Along with these, in the same example, there can be environmentalists, health advocates, and state and local governments who seek different regulatory outcomes. Capture Theory offers no logic for predicting which among many interest groups will capture the politician.

This inherent weakness was addressed when Nobel Laureate Stigler (1971) developed the economic or Special Interest Theory of regulation. Professor Stigler suggested that to predict which interest groups will prevail in a regulatory contest, one should imagine an auction where the politician offers her vote to the highest bidder. The interest group that bids the most will be the one with the most to gain or the most to lose if unable to prevail. Of course, having a low-cost advantage in organizing a winning bid within an interest group involves developing the bid,

dealing with dissenters, and in doing all this, minimizing transaction costs. Stigler's model of enriched capture theory offers a large dose of insight for those who wish to understand regulatory outcomes.

There is yet one more important dimension to consider when elaborating theories of regulation. Those who demand regulation or seek to avoid regulation are willing to support politicians who accommodate them. But while the special interests who seek regulation need no explanation when the promised rules are delivered to them, politicians must justify their actions to their broad support base. They must be prepared to explain their actions in terms that go beyond simply trying to assist an industrial group in gaining monopoly power. Doing so is simplified when the politician can make a moral appeal and state that he was simply trying to do the right thing, that he was attempting to serve the public interest. Combining public interest justification—"We must take care of the environment for the sake of our children!"—with special interest demand for regulation—"Give us a rule that sets higher standards for new entrants than established ones"—introduces Yandle's (1983) Bootlegger/Baptist theory of regulation.

2.3 The Road to the Garden Is Paved with Good Intentions

The Bootlegger/Baptist theory of regulation draws its name from episodes in rural America that involved the regulation of the sale of alcoholic beverages. Historically in rural areas there were two groups that supported state laws that shut down liquor stores on Sundays. These were Baptists, religious groups who opposed the sale of demon rum at any time, but especially on Sunday, and Bootleggers who bought their booze on Saturday and resold it on Sunday, at a profit, when legal competition was eliminated. One group, the Baptists, gave the politicians moral justification for shutting down the legitimate sellers. The Baptists also monitored the situation to make certain the laws were enforced. The Bootleggers, on the other hand, welcomed Baptist enforcement assistance and sometimes provided financial contributions to the political campaigns of local law enforcement officers as well as to elected politicians who made certain the desired laws were renewed regularly.

When Sunday closing laws were up for renewal, the Bootleggers never marched in front of state capitals looking for ways to raise rivals' costs. They did not have to. The Baptists did that for them.

The Bootlegger/Baptist theory requires at least two kinds of interest groups. One provides moral justification for a regulatory action. The other group, which is in it for the money, helps to grease the political rails. It is, of course, possible for there to be multiple Bootlegger and Baptist groups, and we will illustrate this later. Historically, in the classic case, the two groups sought the same outcome and, while never meeting to conspire, were willing to struggle mightily to succeed. At the height of its success, this powerful pairing entirely shuts down the legal sale of alcoholic beverages in counties, states, and—during Prohibition (1920–1933)—the nation as a whole (Boudreaux and Pritchard 1994). We should quickly note,

however, that Bootleggers never supported laws that limited consumption, although their Baptist brethren might. Instead, the two dissimilar groups found middle ground with closing laws that limited the extent of the market. The point is a critically important one that we will emphasize later. The Bootlegger/Baptist theory may not be useful in explaining the fact that regulation limiting the market for alcoholic beverages occurred when it did, but it can be powerful in explaining the fine print of regulations.

The theory helps to explain why, for example, federal laws designed to improve air and water quality impose CAC regulations that are stricter and more costly for new pollution sources than for older ones (Yandle 2013), why costly scrubbers were required for electricity-generating plants no matter which coal—clean or dirty—they burned (Ackerman and Hassler), and why major tobacco companies, state governments, and health advocates jointly oppose the unregulated sale of e-cigarettes (Adler et al. 2013). In a similar way, environmental groups and producers of natural gas and solar panels support regulations that control greenhouse gas emissions, the former due to concerns about global warming, the latter in pursuit of a competitive advantage (Buck and Yandle 2002).

The present campaign to regulate the largest transportation company, Uber, is a further illustration of the working of the Bootleggers, the taxicab drivers, and the Baptists, those concerned with passengers' safety and welfare, which are allegedly compromised in the absence of government regulation. Bootlegger/Baptists coalitions understandably can lead to situations where those in it for profit subsidize those who take the moral high ground and are thus better positioned to become outspoken regulatory advocates. For example, U.S. producers of natural gas made major contributions to the Sierra Club when the environmental organization was lobbying the U.S. Environmental Protection Agency to impose high-cost regulations on coal-fired electricity generating plants (Walsh 2012). In a sense, the Bootleggers became Baptists—or were at least baptized.

We point out again that successful Bootlegger/Baptist support of a particular regulation does not mean that the outcome is bad for the economy or for society as a whole, somehow measured. Transaction politics is part and parcel to the operation of the U.S. political economy. Rent-seeking drives interest groups to organize and then compete in the political commons with other rent-seekers. But again, we note that just as it is possible for an outcome to be good, it is also possible for it to be evil. When Bootleggers and Baptists enter the Garden of Good and Evil, anything can happen.

3 The Dynamics in the Garden of Good and Evil

In our model, Baptists want a targeted activity to stop. They believe consumption of alcoholic beverages, for example, is harmful to both the actor and society. Bootleggers do not want the activity to be stopped; instead they want to corner the market. A Bootlegger would like the activity to be *restricted*—so long as they are

positioned to reap windfall profits from its restriction. The Baptists use persuasion as a means to end the activity. Their arguments resonate because, while they may sometimes be misguided, they are delivered in earnest. The Baptist realizes, though, that for activities for which the demand is very inelastic, there will always be those for whom persuasion alone is never persuasion enough. They will resort to the power of the state only as a last resort. The Bootlegger resorts to the state on principle. Bootleggers are not constrained in their rhetoric to truthful statements, whether delivered in earnest or not. Nor do they feel constrained to the use of rhetoric alone; they are merely interested in what works.

On its face this arrangement would seem unlikely to consummate. If Baptists recognize that the motivation of the Bootleggers is not to end the activity, but instead to profit from it, the Baptists might terminate their lobbying activity. If the Bootleggers believe that their Baptist partners will not stop until the activity, and hence the Bootleggers' profits, are eliminated entirely, they might part ways. Yet the Garden seems to be a place where these details may be overlooked. Why? It is because many Baptists are realistic. They are willing to settle for a reduction of the unwanted activity because they recognize that a full cessation is unlikely through persuasion alone. Many Bootleggers recognize this and are thus willing to collaborate with the Baptists toward a full cessation that they feel confident will never come.

Unpopular restrictions on desirable human activities are difficult to enact and to enforce. By definition, it is almost impossible in the first place to convince people to willingly give up activities that they highly value (whether in the psychic or material sense),[2] let alone to support legislation restricting them. Enforced compliance is both expensive and obnoxious. Building the apparatus of detection and punishment soaks up resources, invites disobedience, and delegitimizes the enforcers. But a persuasive argument can not only increase voluntary compliance—thus obviating the need for some enforcement—but may even in the best case produce propagandizing and self-policing efforts by parts of the population, thus reducing compliance costs further. In other words, persuasive rhetorical arguments are inexpensive and they emit positive externalities: the reduction in the cost and the increase in the effectiveness of expensive and risky coercion. Together the Bootleggers and Baptists are more effective than alone.

So while they often work separately—though with unconscious parallelism—Bootleggers and Baptists can find a way to work together despite the obvious lack of meeting of the minds. They do so by creating a positive sum two-party game that capitalizes on their joint value. Simply put, the Baptist's rhetoric[3] combined with

[2]People love to drink even though it may offer no material profit; people love to cash checks from their ownership stake in coal-powered electric plants, even though it may offer no psychic benefit.

[3]Baptists may be said to use threats/coercion, but not directly. They would refer to eternal damnation, or to catastrophic damage to the environment, or depletion of resources, etc., but these are always external theoretical threats and come as the result of having performed the activity itself, not threats of arbitrary force here and now with the purpose of preventing the activity from taking place.

the Bootlegger's pragmatic manipulation of the political process makes legislation favorable to their joint cause more likely to be passed, less expensive to maintain, and profitable to both.

3.1 The Logic of Political Survival

In their influential works *The Logic of Political Survival* (2004) and *The Dictator's Handbook* (2012), political scientists Bruce Bueno de Mesquita and Alastair Smith developed a framework for understanding political choices and welfare outcomes in different polities. We selectively adopt and adapt their framework to examine the case of Bootleggers and Baptists political involvement in CAC regulation in the U. S.

According to the authors, attaining power is just the first step for the aspiring leader. Achieving the throne does not in itself confer the power to do whatever the new leader wants. Political leaders, from autocratic dictators to democratically elected presidents, must act to survive in the political environment. Once elected, the leader's actions are determined first and foremost by his political environment and *not* by his civic-mindedness. In politics, altruism is a luxury good dearly bought.

In any polity there are people who are authorized to take part or advocate in the political process, and from whom at least tacit consent to political outcomes is required. This is the *selectorate*. These are the *interchangeables* (2012). They do not matter individually, only that they peacefully participate as a group in whatever capacity they have been granted. They allow, for example, for there to be a ruling family or for competing candidates on a ballot.

The *influentials* are the subset of the interchangeables whose consent to a particular candidate is necessary for that candidate to proceed to inauguration (2012). The influentials matter individually to the extent that their vote is counted or their support is recognized, but they are replaceable from the ranks of the interchangeables if the price is right. They decide the winner of the competition that the interchangeables have allowed. The influentials are the victorious group, formerly mere interchangeables, whose preference for a particular candidate, family, or other ruling entity has prevailed. Naturally they will expect something in return. The losing group of interchangeables will hope instead that the cost of their defeat is not too onerous.

Finally, there is the subset of individuals from whom aid and support is absolutely necessary to gaining and maintaining power, the *essentials* (2012). This can be a small subset of the influentials actually responsible for selecting acceptable candidates. Before you can win a competition you have to be in it, and the essentials get candidates onto the ticket in the first place. Their continued support is crucial to the long-term survival of the leader and they are expensive and difficult to replace.

Generally speaking, in a democracy, all eligible voters comprise the interchangeables. Their general consent matters but they are fungible. The influentials

are the subset of the selectorate whose support will decide the contest (e.g., 50 % + 1 in the American system). Their turnout on polling day matters, and it comes with expectations. The essentials, or the winning coalition, are the even smaller number of critically important individuals, whether financial backers, political power brokers, or trusted confidants. To achieve and to maintain power, all leaders will have to devise a political strategy to satisfy each of these groups in return and in proportion to their usefulness. It is the relative size of the three groups that influences and helps to define the political strategy used to capture power, and the resulting political outcome.

3.2 Public and Private Goods

According to Bueno de Mesquita and Smith, the leader can choose to mix a bundle of two types of goods, Public and Private, to induce loyalty and support. Private goods are goods which can be enjoyed directly by individuals or small interest groups. These include a wide range from cash to privileges: a car, an Ambassadorship, or a commutation of a punishment. Private goods come with both positive and negative externalities. They are targeted and effective: valuable to recipients with very little wasted on those for whom receipt is neither necessary nor intended. This in turn encourages influentials to compete for the essential status. They are, however, expensive on a per recipient basis, and can induce resentment, especially in those who foresee no chance on the horizon of enjoying the benefits. Benefits, they remember, that were financed with their tax dollars.

Public goods are goods enjoyed by large portions of the polity, if not the nation at large. They might include things from local schools and monuments to interstate highways and the rule of law. Similar to Private goods, Public goods also contribute both costs and benefits to the leader's calculation. On the one hand, they are expensive and not well targeted. Very often, essentials, influentials, and interchangeables alike enjoy these goods and services. But this is not entirely a defect. Such externalities enable the Public goods to be cost effective for large groups and are a palliative to the losing coalitions.

Because of these different and complementary features of Private and Public goods, the bundle that the leader distributes to society will contain a mixture of both. The ratio of Private to Public goods in the bundle will be defined by the relative distribution of the three groups, the essentials, influentials, and interchangeables, in the polity.

For example, in the U.S., hundreds of individuals and special interest groups are essential to finance the campaign, and tens of millions of votes are needed on election day to influence the decision, while the patient consent to temporary defeat is required of tens of millions more. Dictatorships are instead defined by the small number of essentials and influentials relative to that of interchangeables. When there are only a few members who must be satisfied with individual gifts, as in the case of a dictatorship, a leader must lean heavily on Private goods. As the number

of these invested parties increases, e.g., in democracies, Private gifts become costly to apportion and thus the leader increases provision of public goods.

Bueno de Mesquita and Smith conclude that the prosperity generally enjoyed in democratic societies is a result of a larger ratio of spending on Public relative to Private goods; that this ratio is determined by the relative representation of key groups in the polity; and that all of this is put into practice by a leader constrained to do exactly that. In other words, it is not from the benevolence of leaders that We the People should expect our public goods, "but from their regard to their own interest" ([1776] Smith 1982).

And leaders are nothing if not self-interested. So for democracy it seemed to work out well enough for a while. But as Mencken famously put it: "Democracy is the theory that the American people know what they want, and deserve to get it good and hard" ([1916] Mencken 2009). The prosperity that came with the abundance of Public goods began to bear suspicious hybrid fruits when crossed with Private goods in the Garden. Just as a synergy between the qualities of Bootleggers and Baptists led to an unholy matrimony, a hybrid good taking on qualities of both Public and Private goods was soon cultivated.

3.3 The Hybrid Good

The framework developed by Bueno de Mesquita and Smith is very insightful. We add to their dichotomy of goods—Private and Public—to explain some essential aspects of strategic political behavior in modern American democracy. Specifically, we add to their framework the fruit of the Garden of Good and Evil—command-and-control regulations (CAC), the Hybrid good.

A CAC regulation contains the features optimal for a democratic leader seeking political support. It can take on the positive features of both a Private good and a Public good and, in some contexts, even minimize their negative properties from the perspective of the leader. For instance, the benefits of a CAC regulation can be targeted to some particular interest groups, and cause outsiders to compete for them as though they were private gifts. It can also provide to the leader the benefits of a Public good by conveying real or perceived benefits to a large number of citizens from across the political spectrum at a low per capita cost.

Unlike most traditional Public goods, CAC regulations are similar to the rule of law in that they do not have to be built out of the public purse. Building a dam is expensive; writing legislation is relatively cheap. The real cost arises after implementation and is usually too opaque and dispersed to be recognizable. Because CAC regulations are very complex, it may be difficult to know who is a winner and who is a loser, and what the ultimate ramifications of the policy will turn out to be. Very interested parties, the types who like to make themselves essential to the process, will not only understand the ramifications but will help to craft them. Rationally ignorant voters, on the other hand, whether influentials or interchangeables, will likely base their opinions on the carefully crafted rhetoric

used to advertise the complex regulation, rather than on its real consequences. Thus, much of the cost of both Private and Public goods is eliminated while retaining many of the benefits of each.

Command-and-control regulation can be targeted to punish groups all the while carrying both the veneer of public spiritedness and an opacity that renders them almost immune to attack. For example, revised emission standards that place severe limitations on carbon dioxide discharge carry a heavy burden for electricity producers with lots of coal-fired plants, but place little burden on all producers that employ gas-fired and nuclear power plants. Buyers of electricity in the first case will face significantly higher prices, but will not likely be able to link the higher prices systematically to newly imposed emission standards. Meanwhile electricity consumers in unaffected regions can happily go forward without experiencing any inconvenience.

A command-and-control regulation may generate appropriable rents for those who benefit from its direct effects. For example, an air pollution control technology newly mandated for all firms in an industry imposes no cost on members of the industry that already use the technology. At the same time, the higher cost imposed on others limits industry expansions and leads to higher prices. All the while, interest groups that value cleaner air will celebrate the new rules even though it is impossible for them to measure changes in the amount of cleaner air that might be produced. Clean air lovers will encourage the regulators while often condemning industrial firms that continue to operate in the regulated environment. Of course, those firms that gain profits from the rule will celebrate too.

Not everyone will celebrate, however. Those who manage to recognize that they are hurt by the regulation, those who suffer the negative externalities such as higher prices or diminished opportunities and are able to make the connection to the CAC regulation in question, may be encouraged to join a coalition working against the rules that are imposing cost on them. But as long as the opportunity cost of collective action (either lobbying against the regulation or for an offsetting one) exceeds the cost imposed by the perceived harm, the harmed individuals will not act to change the status quo (McCormick and Tollison 1981). It is only harm in excess of the cost to avoid the harm that drives victims to defensive action.

Unlike the budget impacts of public works projects, which will have a listing of itemized costs in a government budget, regulations are born without clearly communicated cost estimates. The great advantage of the regulatory transfer mechanism is that it is difficult for all but those directly involved to truly understand its costs and effects. Indeed, regulations can become so complex that some who believe they are beneficiaries may only learn later that they have been hoodwinked. Consider for example, the U.S. requirement that ethanol be blended in all gasoline for allegedly environmental reasons. The ethanol blend diverts corn production away from food to energy, raising the price of both. Meanwhile, evidence accumulates indicating that burning ethanol blends are more harmful to the environment than unblended gasoline.

In spite of situations like ethanol, such regulations come with the claim that they are for a good cause, and thus their dual identity makes them an attractive choice for

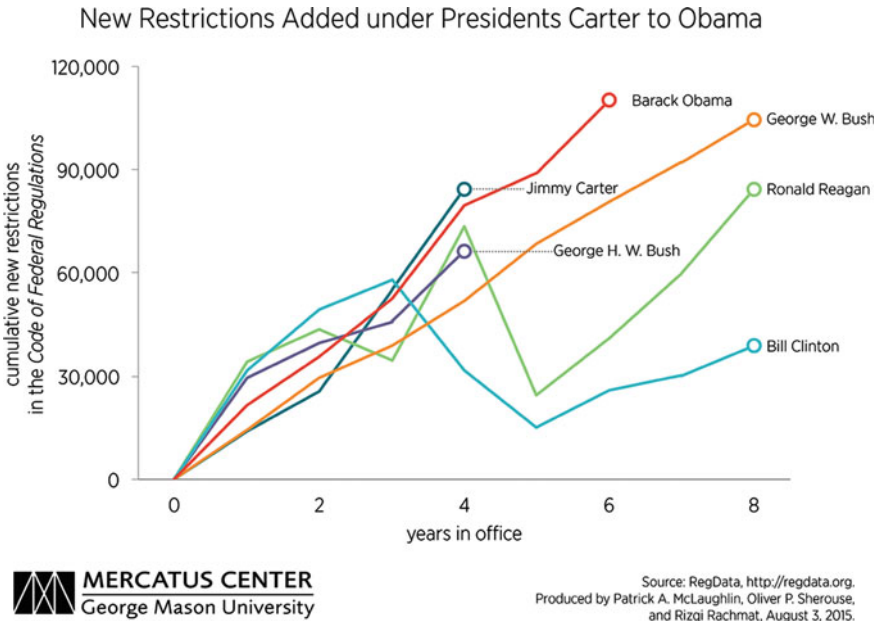


Fig. 4 The increase in the use of command-and-control regulations across presidents: from Carter to Obama

politicians who seek to garner political support while minimizing the political cost of their choice. Because of this dynamic, CAC regulations have taken root like weeds in the Garden of Good and Evil.

What about the extent of all this CAC activity? How burdensome is it? We know of no way to accurately estimate the economic cost of all CAC regulation, but we can offer a proxy for the growth of CAC by drawing on data produced by George Mason University's Mercatus Center. The Mercatus Center has developed a regulation severity index that is based on an annual count, by industry, of the frequency of command-and-control words found in the U.S. Code of Federal Regulation, which is the final compendium of all current federal rules. Their RegData index counts the frequency of the words Must, Shall, May Not, Prohibited, and Required.[4] Figure 4 reports the results for six U.S. presidential administrations that include the first 6 years of the Obama administration. We note that Carter, Clinton, and Obama were Democrats; George H.W. Bush, George W. Bush, and Ronald Reagan were Republicans.

The data demonstrate that Republican and Democratic administrations alike, from President Carter to Obama, have continually added more and more CAC regulatory language to the books. Effectively designed, command-and-control

[4]RegData is available at http://regdata.org.

social regulations can take this hybrid form that silently grants monopoly rents to some firms and industries while simultaneously, and loudly, providing a public good with the moral high ground. In other words, the successful CAC regulation tells the rent-seekers what they want to hear and what it wants the public to overhear.[5]

We point out that in the Garden of Good and Evil, a perfect regulation may be seen as one that maximizes benefits to the elected leader's essentials, advertises benefits to nonessentials, and maintains sufficiently disbursed or opaque costs to avoid resistance and effective rhetoric to reduce enforcement costs. Designing such subtle rules is by no means an easy task. Dressing private gifts in the cloak of public mindedness requires a team of expert tailors. The team consists of both the Baptists and the Bootleggers working to influence the regulation. Neither the Baptists nor the Bootleggers can achieve this effective regulatory provision when working alone.

3.4 The Bootleggers as Rent-Seekers in Consumer Markets

Recall that in our earlier discussion we described Bootleggers first as illicit sellers of alcoholic beverages in situations where state or federal law shut down the sale of those beverages on particular days of the week or entirely. The focus there was on a consumer good. In that situation, the Bootlegger is one who demands, in an economic sense, a restriction on the good—in order to become the monopoly supplier on either the legal or the black market. The more inelastic the demand is for the good, the more likely that people will continue its consumption despite the increased price or possible legal ramifications. That inelasticity of demand was an essential feature of what constituted a traditional Bootlegger zone of activity.

The U.S. regulation of tobacco illustrates the point (Yandle et al. 2008). Early federal regulation limited cigarette advertising and marketing practices, an action that enjoyed strong support from healthcare advocates *and* tobacco companies—*with already highly developed brands and market share*.

Still, while the Bootleggers and Baptists celebrated, no doubt in private for the Bootleggers, legislation arose that called for equal TV time for public health messages when tobacco companies advertised. As a result, major tobacco companies saw their sales fall. They were horrified. But then, the regulators called for a ban on all cigarette TV advertising. Healthcare advocates celebrated along with the Bootleggers once again as their profits rose in the face of hamstrung new competitors.

[5]This conclusion is adapted from a line by Anthony de Jasay (1985, 7) in *The State*: "It seems to me almost incontrovertible that the prescriptive content of any dominant ideology coincides with the interest of the state rather than, as in Marxist theory, with that of the ruling class. In other words, the dominant ideology is one that, broadly speaking, tells the state what it wants to hear, but more importantly what it wants its subjects to overhear."

Consider now the case of marijuana prohibition, another consumer good for which there is a high inelasticity of demand. Marijuana growers and sellers have an incentive to keep marijuana illegal so that they can continue to reap the profits from its sale on the black market. Law enforcement and incarceration professionals prefer to sustain both the illegality and the use of marijuana—and thus a steady supply of "customers"—in order to protect their jobs. These are traditional Bootleggers. They prefer the activity ongoing but restricted, so that they can profit from the trade.

Rent-seeking Bootleggers, on the other hand, seek to profit off of a total prohibition by the sale of some substitute. Some representatives from the plastics industry, for example, keen on preventing the use of stronger and more durable industrial hemp fibers, have joined in the effort to ban marijuana. Additionally, producers of painkillers and nausea-reducing pills may support a ban in order to eliminate it as a substitute for their pharmaceuticals.

Here, in the case of marijuana prohibition, we have identified groups representing two types of Bootleggers: illegal producers and prison or law enforcement agents, representing traditional Bootleggers, and pharmaceutical producers and plastic groups representing rent-seeking Bootleggers. The Bootleggers cannot expect to prevail alone when lobbying for continued restriction of the production and sale of marijuana. Their hope comes in the form of Baptists with the moral high ground health, safety, and the common weal. If a majority can be convinced that it is harmful, dangerous, and tears the fabric of society, then they will consent and assist in, or at least tolerate, enforcement.

3.5 The Baptists: Idealists and Realists

With respect to the historical case of Prohibition in the U.S., our use of the term Baptists is meant to describe the type of people who struggled for the prohibition of alcohol on religious, moral, or ethical grounds. The term describes the true believers, those who believed that alcohol corrupted the souls of men and imposed high cost on American family life. In general, the Baptists power lies in their ability to "move men's emotions" (Bongiorno 1930, 358) in a way that they become willing to change their behavior in alliance with a certain code of action—regulation—that satisfies the Baptist view of how the world should be organized.

The word Baptist in this sense has now become a metaphor for this brand of regulatory rhetoric, that is, the advocacy for the use of state power to enforce a prohibition or restriction on a good or activity in the name of justice or decency or social wealth maximization. We note that Baptist arguments for setting limits on an activity do not necessarily rely on scientific or logical justification. Indeed, Baptist groups are more inclined to argue that moral arguments trump scientific ones.

For instance, Baptist groups who oppose the use of any synthetic chemicals in the production of food may base their arguments on false assumptions, e.g., that the use of any dose of any synthetic chemical is equally toxic, that synthetic chemicals are more dangerous than natural chemicals, etc., even in the face of credible

contrary evidence. The Baptist's theories may not be scientifically sound, but logic and soundness are not the point, not if we are trying to understand how regulations come about and last on the books for decades in the real world. As Andrew Bongiorno correctly pointed out: "Theories of this kind may be absurd from the scientific point of view, yet a social scientist who views them not as absurd verbiage, but as social facts, cannot but recognize them as powerful determinants of the social equilibrium" (Bongiorno 1930, 358).

On a spectrum of regulatory stringency, we can divide Baptists into two types that deserve special scrutiny: those who desire absolute prohibition (i.e., maximal regulatory stringency) and those who will settle for some level of decrease in the quantity demanded of the particular activity. Let us call the first type the idealists and the second the realists. The Anti-Saloon League during prohibition is an example of the idealist Baptists. On the other hand, environmental groups worried about anthropogenic global warming who seek the reduction in carbon emissions but not its elimination are an example of the realists. The latter are satisfied with winning on the margin.

For example, a realist Baptist who would like to prohibit the use Bisphenol-A (BPA) in all plastic products would be willing to work with a rent-seeking Bootlegger industrialist who wishes to boost the sale of his Bisphenol-S plastic products by prohibiting BPA. While the realist Baptist's desire for banning BPA is grounded in concerns over health and safety, he does not require the same ethical standards from the industrialist, who he may realize is merely interested in raising a rival's cost. To a Baptist realist the end justifies the means.

The same is not true for an idealist, however. An idealist may only support the industrialist's campaign if the industrialist's rhetoric is a true reflection of his intentions. For example, an idealist Baptist might lobby for CAC regulation in the food industry along with a health food grocer who also sincerely believes that nonorganic foods are a danger to public health—and stands to profit handsomely from competition hampered by the proposed regulation. In this case, the boost to profits in the grocer's sector would be a happy coincidence of action based on sincerely held convictions and a synergy of method.

Neither the idealist nor the realist, however, would *knowingly* join a traditional Bootlegger. Recall that a traditional Bootlegger is one who is not only hiding his true intentions (he has no moral qualm over the product at issue), but whose goal is ultimately to profit from the sale on the black market of BPA produced goods. Therefore, a necessary condition for coalescence of Baptists and the traditional Bootleggers, or Baptists and the rent-seeking Bootleggers, is agreement on the rhetoric for their policy position. In other words, both the traditional and the rent-seeking Bootleggers must publicly support the Baptists' rhetoric if they desire the union to succeed.

Baptists can still achieve their goal by working with rent-seeking Bootleggers, even if at the expense of their preference for sincerity. Tragically, however,

whenever the Baptists seek the same outcome as traditional Bootleggers they move further from their beloved goal and achieve the opposite of their intention. And in the world of transactional politics the result is always more pages in the *Federal Register*.

3.6 Collaboration as a Cost-Minimizing Technique

With political leaders eager to shower supporters with gifts in exchange for office, Baptist powers of persuasion and pragmatic Bootlegger power to produce regulatory copy provide a complete toolbox. We are left to show how their work reduces the production and maintenance costs of regulations.

The "profits" derived by consumers from conforming to the Baptist's argument are ethical and moral in nature; they are not tangible or immediate like the benefits perceived from, say, taking a long, hot, guiltless shower. The nature of the profits advertised by Baptists to consumers means then that the rewards are extremely valuable to those who are convinced that they exist, yet worthless to nonbelievers. And even when morality is symmetric across individuals, their discount rate might differ dramatically. This is the heart of the reason why Baptists can never achieve full compliance on their own. They cannot *prove* to people, but instead must strive with great difficulty to convince them, that the value of their future rewards is worth the required present sacrifice.

The prospective profits of the traditional and rent-seeking Bootleggers, on the other hand, are readily convertible to cold, hard cash. This form of profit is perfect for greasing the wheels of favored legislation by filling the coffers of favored legislators. The grease being necessary since the Bootlegger obviously cannot advertise the personal profits and social costs associated with it—common understanding would either deliver the product stillborn or retard compliance and thus profit margins. The Bootlegger therefore relies on the political leader to enforce compliance in exchange.

The political leader also recognizes costs and risks involved with enforcing unpopular prohibitions. Enforcement without complementary rhetoric is expensive and labor intensive; in isolation the optimal level of compliance is not 100 %. And the power of the sword breeds resentment. Given the leaders' desire to maintain office, he will not consent to such a deal; it is political suicide.

The presence, and sometimes the visible collaboration, of the Bootleggers and Baptists is necessary before an onerous CAC regulation can be passed, let alone consistently and efficiently enforced. Therefore, the leader is much safer to act after the Bootleggers and Baptists have formed a symbiotic coalition. The Baptists will convince the selectorate of "the ethics" of the regulation, and the Bootleggers will work to enforce it upon those who remain unconvinced. The environment has sent consistent feedback: successful attempts will be those that maximize compliance and minimize cost of enforcement.

4 Final Thoughts

In closing, we wish to make three points. First, regulations are vehicles for transferring resources by the way of transactional politics. To regulate is to assert a form of property right over a resource, or to rearrange existing property rights, or to create rights in previously "unowned" areas. In order for property rights to serve their purpose, they must be secure. So in order to gain the political support of the essentials and influentials, the aspiring political leader when seeking office must convince these vital parties that his offer is credible and will be secure against future repeal. Credibility rises with effective Baptist rhetoric in support of the Bootlegger enterprises, whether in unwitting or explicit coalition. Low Baptist engagement leads to low return on the political capital invested by Bootlegging interest groups in their effort to elect and maintain a friendly political leader. And, of course, if the Baptists leave entirely, for whatever reason, then the exposed Bootleggers lose their rhetorical cover and political footing.

Second, drawing on Public Choice logic and Bootlegger/Baptist theory, our chapter has offered an explanation as to why CAC regulation dominates the U.S. regulatory landscape. Critical to our story are hybrid regulations, those that provide appropriable private rents for Bootlegger interest groups and public good benefits for the Baptist influenced constituencies. Hybrid regulations reduce political transaction costs for any politician hoping to maintain political power. Our presentation of data on U.S. regulatory activity demonstrates that U.S. regulatory activity goes on unabated irrespective of the political party in power.

As a final point, what can we say about outcomes in the Garden of Good and Evil? Is there any evidence that CAC regulations generate more normatively evil than good or vice versa? Bueno de Mesquita and Smith predicted a positive monotonic relationship between economic activity and the more democratic a polity is as defined by the relative makeup of the essentials, influentials, and interchangeables (de Mesquita et al. 2004). The relationship seems to bear out—to a point. When adopting and adapting their framework to include hybrid regulatory provision and a Bootlegger and Baptist coalition we are able to better comprehend the growth of CAC regulation in the present incarnation of democracy in the U.S. and its potential deleterious effects. With the turn of every administration, different Bootlegger coalitions will get their turn to pull the lever.

It also is extremely difficult to provide a positive answer to the net benefit question because at the heart of weighing the good and the evil is an assumption of *who's* good and *who's* evil? Who is worth counting in the social calculus, and at what rate? Nevertheless, we will present two arguments, one empirical and the other theoretical, in an attempt to understand the outcome's characteristics.

The first argument is Davies (2014a), who built on the RegData measure of regulatory stringency to study the effect of CAC regulation on the productivity of highly regulated enterprises and industries relative to those that carry a lessor

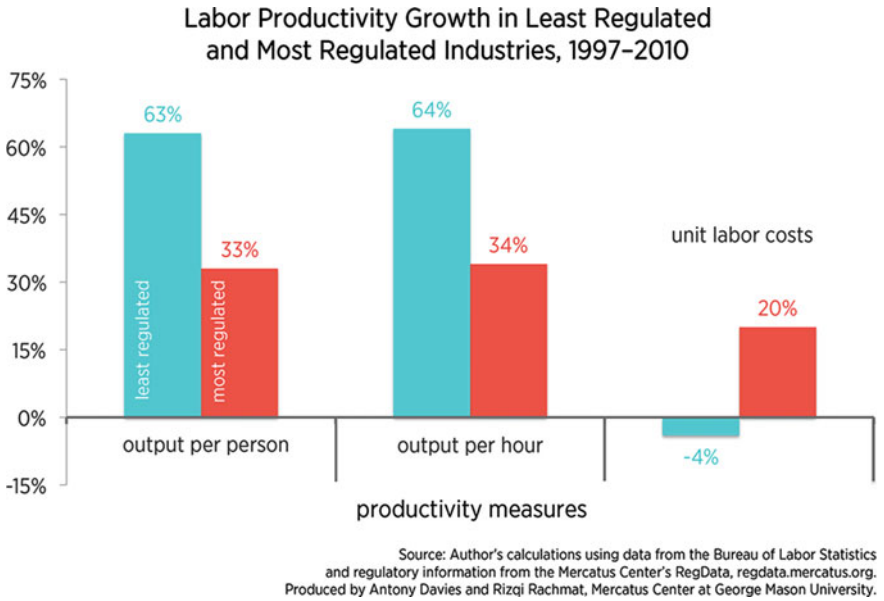


Fig. 5 Growth in labor productivity and cost in least and most regulated industries: 1997–2010

regulatory burden (Davis 2014b, 11). Davies relied on the Bureau of Labor Statistics measure of production efficiency for 51 industries over the span of 14 years. He calculated relative regulatory stringency for each of the industries and then broke his sample of 51 industries into three equal parts, those heavily regulated, moderately regulated, and lightly regulated.

A ray of insight into this matter is provided in Fig. 5. When comparing growth in output per person, output per hour, and unit labor costs for two samples of U.S. industries—those most heavily regulated and those regulated least—across the years 1997–2010, we can infer that more frequent CAC regulation entangles production, leads to lower levels of output per worker and lower growth in income.

The second argument is that of Jane Jacobs's as advanced in her seminal publication *Systems of Survival: A Dialogue on the Moral Foundations of Commerce and Politics* (1993). Jacobs did not conduct an empirical analysis of welfare outcome; instead she studied the *processes* by which different outcomes emerge. She did so to infer the efficacy of outcomes from the processes generating them.

To Jacobs, there are two types of syndromes underlying the working of any society: one syndrome she names the "guardian syndrome" and the other she names "the commerce syndrome." The guardian syndrome is based on the "taking" strategy of survival, a characteristic of the government sector and of Bootlegger methods. The commerce syndrome is based on the "trading" strategy of survival, a characteristic of the private, or voluntary, sector, and of Baptist methods. For any society to prosper both syndromes may be necessary, but it is crucial for them to

remain separate. When the private sector, which is concerned with generating profits, trades its support of the government sector for exclusive grants of privilege, we get what Jacobs terms "monstrous moral hybrids," which can yield stagnation by regulation.

References

Ackerman, Bruce, and William T. Hassler. 2008. *Clean Coal, Dirty Air*. Cambridge: Harvard University Press.

Adler, Jonathan, Andrew Morriss, Roger Meiners, et al. 2013. Bootleggers, Baptists and E-Cigs. *Regulation.* 38(Spring): 30–35.

Bernstein, Marver. 1977. *Regulation by Independent Commission*. Princeton: Princeton University Press.

Bongiorno, A. 1930. A study of Pareto's treatise on general sociology. *American Journal of Sociology*: 349–370.

Boudreaux, Donald J., and A.C. Pritchard. 1994. The Price of Prohibition. *Arizona Law Review* 36: 1–10.

Brito, Jerry and Susan E. Dudley. 2012. *Regulation: A Primer*, 2nd edition. Fairfax, VA: Mercatus Center at George Mason University.

Buck, Stuart, and Bruce Yandle. 2002. Bootleggers, Baptists, and the Global Warming Battle. *Harvard Environmental Law Journal* 1(26): 177–229.

Bueno de Mesquita, Bruce, and Alastair Smith. 2012. *The Dictator's Handbook: Bad Behavior Is Almost Always Good Politics*. New York and London: Public Affairs.

de Mesquita, Bueno, Alastair Smith Bruce, Randolph Siverson, et al. 2004. *The Logic of Political Survival*. Massachusetts: The MIT Press.

Davies, Antony. 2014a. More-Regulated Industries Experience Lower Productivity Growth. *Mercatus Center at George Mason University*. http://mercatus.org/publication/more-regulated-industries-experience-lower-productivity-growth.

Davies, Antony. 2014b. Regulation and Productivity. *Mercatus Center at George Mason University.* http://mercatus.org/sites/default/files/Davies_Regulation&Productivity_v1_0.pdf.

de Jasay, Anthony. 1998. *The State*. Indianapolis, IN: Liberty Fund.

Jacobs, Jane. 1993. *Systems of Survival: A Dialogue on the Moral Foundations of Commerce and Politics*. New York: Random House.

Kolko, Gabriel. 1963. *The Triumph of Conservatism: A Reinterpretation of American History: 1900–1916*. Chicago, Ill: Quadrangle Books.

McCormick, R. E, and R. D Tollison. 1981. *Politicians, Legislation, and the Economy: An Inquiry into the Interest-Group Theory of Government*. Boston; Hingham, Mass. In *Distributors for North America*, ed. Nijhoff, M. Boston: Kluwer.

Mencken, H.L. 2009. *Little Book in C Major*. South Carolina: Biblio Bazaar.

Pigou, Arthur C. 1920. *Economics of Welfare*. London: Macmillan and Company.

Shamoun, Dima Y. 2013. *The Propensity to Truck, Barter, and [impede] Exchange: Democracy the Unknown Ordeal* (Doctoral Dissertation, George Mason University). Retrieved from http://digilib.gmu.edu/xmlui/bitstream/handle/1920/8756/Shamoun_gmu_0883E_10516.pdf?sequence=1&isAllowed=y.

Smith, Adam. 1982. *An Inquiry Into THE Nature and Causes of the Wealth of Natilons*. Indianapolis: Liberty Fund.

Smith, Adam, Richard Wagner, and Bruce Yandle. 2011. A Theory of Entangled Political Economy, with Application to TARP and NRA. *Public Choice* 148: 45–66.

Stigler, George J. 1971. The Economic Theory of Regulation. *Bell Journal of Economics & Management Science.* 1: 3–21.

Wagner, Richard. 2009. Property, state, and entangled political economy. In *Markets and politics: Insights from a political economy perspective*, eds. Wolf Schäfer, Andrea Schneider, and Tobias Thomas, 37–49. Marburg: Metropolis.

Walsh, Bryan. 2012. How the Sierra Club Took Millions from the Natural Gas Industry—and Why They Stopped. *Time*, February 2. http://science.time.com/2012/02/02/exclusive-how-the-sierra-club-took-millions-from-the-natural-gas-industry-and-why-they-stopped/.

Yandle, Bruce. 1983. Bootleggers and Baptists: The Education of a Regulatory Economist. *Regulation*, 12–16.

Yandle, Bruce. 2010. National TV Broadcasting and the Rise and Decline of the Regulatory State. *Public Choice* 142(March): 339–353.

Yandle, Bruce. 2013. How Earth Day Triggered Environmental Rent Seeking. *The Independent Review* 1(18): 35–47 (Summer 2013).

Yandle, Bruce, Andrew P. Morriss, Andrew Dorchak, and Joe Rotondi. 2008. Bootleggers, Baptists and Televangelists. *University of Illinois Law Review* 4: 1225–1284.

Hobbesian and Contractarian Constitutions

Geoffrey Brennan and Giuseppe Eusepi

1 Contractarian Methods and Constitutional Analysis

Within the constitutional political economy tradition, the predominant normative framework is contractarian. That is, the standard test in evaluating political institutions is whether such institutions could conceivably [or more rarely, actually did] emerge from free agreement among rational individuals. Clearly, the application of this test will depend on a range of details: whether the agreement in question is actual or 'conceptual'; what exactly it means to say that such agreement is "free"; what rationality requires and whether the requirement of rationality seriously restricts the possible institutional arrangements that might emerge from actual or hypothetical agreement; and whether the contract in question must be between individuals or might be between groups of some kind. However these issues are resolved, the central role of agreement and its analogical connection to market exchange are central elements in standard "constitutional political economy" (CPE henceforth).

This strong contractarian flavour reflects the influence of Buchanan on the CPE enterprise.

G. Brennan (✉)
School of Philosophy, ANU College of Arts and Social Sciences,
The Australian National University, Canberra ACT 2601, Australia
e-mail: Geoffrey.Brennan@anu.edu.au

G. Brennan
UNC-Chapel Hill, Chapel Hill, USA

G. Brennan
Duke University, Durham, USA

G. Eusepi
Faculty of Economics, Department of Law and Economics of Productive Activities,
Sapienza University of Rome, Rome, Italy
e-mail: Giuseppe.Eusepi@uniroma1.it

A. Marciano and G.B. Ramello (eds.), *Law and Economics in Europe and the U.S.*,
The European Heritage in Economics and the Social Sciences,
DOI 10.1007/978-3-319-47471-7_4

And many of the positions taken in CPE analysis correspondingly reflect Buchanan's views on the "details" already enumerated. So, for example, Buchanan is determinedly individualistic; and reckons 'rationality' to be a distinctive property of individuals—and only contingently a property of groups, for well-known prisoner's dilemma reasons. Accordingly, only agreements among *individuals* can have any ultimate normative authority. And Buchanan is inclined to understand "rationality" 'thinly', in the manner characteristic of ordinary economics. That is, rationality is understood to require that agents act to promote their ends—but the ends in question can be [almost] anything at all, provided they are not blatantly self-contradictory. In addition, the reference to market exchange as a benchmark notion in understanding the notion of agreement reflects a particular conception of agreement—a conception that arises naturally among economists but is less common among political and moral philosophers. In short, the 'contractarian approach' followed in most 'constitutional political economy' is a distinctive one and differs in a number of ways from other kinds of contractarianism—Rawls's, Gauthier's and Scanlon's for example—that inhabit the political and moral philosophy literature.

Moreover, it is clear that normative constitutional analysis might be pursued from other, non-contractarian, normative positions. There seems to be no particular reason why contractarians as such should have a monopoly on concern with institutional arrangements. Alternative sets of political economic 'rules' might be evaluated from utilitarian or neo-republican or libertarian or egalitarian positions; and indeed political theory abounds with such exercises. Constitutionalism does not necessarily involve contractarianism.

There is, however, a suggestion in the Buchanan formulation that the obverse logical relation might apply—that is, that contractarianism implies constitutionalism. Much of the Buchanan interest in constitutional rules arises through an apparent conviction that a shift to the constitutional level of analysis is necessary if contractarian modes of thinking are to have any substantive content. In *The Calculus of Consent*,[1] for example, the explicit reason for shifting to the analysis of decision rules is that the contractarian ideal of unanimity is infeasible [or at least 'too costly'] at the in-period level. The intimate connection between unanimity and the Pareto criterion of the "new welfare economics" is a recurring theme in Buchanan's writings.[2] So is the notion that informational limitations imply that actual exchange (or actual agreement) is the only authoritative test of mutual benefit. Consider:

[1]Buchanan and Tullock (1962).

[2]See, for example, "The Relevance of Pareto Optimality"; "Positive Economics, Welfare Economics and Political Economy" in Buchanan (1999).

> … Unless the observing economist is assumed to be omniscient, his classification of a final position as non-optimal can never be more than a conjectural hypothesis that is impossible to test. If members of the group do not explicitly choose among final positions in the appropriately defined welfare space, the hypothesis that some members of the group can be better off by a change remains empty. … By contrast, the classification of an organisational rule as non-optimal can be considered as a hypothesis that is subject to conceptual testing. If a presumed or apparent non-optimal rule cannot be changed through agreement among members of the group, the hypothesis stating that the rule is non-optimal is effectively refuted.[3]

The idea here seems to be that by moving to the level of choice among rules the scope for agreement is enhanced, and this is both because the informational requirements are less extreme at the constitutional level and because it is not necessary that 'optimal rules' always produce 'optimal results'. What is clear, however, is that if there should be a set of political arrangements which would leave all citizens better off than in the status quo—a set, that is, such that a move from the status quo arrangements to the alternative would command agreement among all [rational] citizens and can indeed be shown to do so—then these political arrangements should be implemented.

It is worth noting that the whole Buchanan approach is framed against several feasibility considerations. One of these is access to relevant information; hence, the reference to the economist's putative "omniscience". Here Buchanan is gesturing towards the abandonment of utilitarian methods under the onslaught of Robbins and Hayek—a gesture that sits well with his own expressed sympathies for subjectivism.[4] Actual exchange is capable of revealing the values that parties place on options available; in the absence of such exchange, the economist has no privileged access to information about values and indeed it is not entirely clear whether such value can be said to exist in the absence of agents' choice behaviour.[5] Another 'feasibility' consideration that has played a critical role in the development of public choice theory specifically concerns assumptions about human motivations—particularly in relation to agents in their political roles. The "benevolent despot" construction has been a focus for public choice critique from the very founding moments of the discipline. The idea that policy advisors and/or their politician/bureaucrat advisees can be appropriately modelled as motivated exclusively to pursue the public interest (however, exactly discerned) has struck a succession of public choice theorists as hopelessly "romantic"—hence, the description of public choice theory as "politics without romance". This assumption, no less than the idea that political considerations are irrelevant in policy choice—the 'despot' aspect of the 'benevolent despot'—is something that public choice expressly aims to set aside. If we focus on agents as they are, if we treat human imperfection as a basic constraint, then we must design institutions that accommodate such imperfection as best we can.

---

[3]Buchanan (1999), p. 228.

[4]See in particular his *Cost and Choice* (1969) and "The Domain of Subjective economics" (1982).

[5]See Buchanan (1991).

In this motivational respect, constitutional choice also offers some relief. In a way similar to that under Rawls' veil of ignorance argument, natural selfishness gives way to a measure of (rational) general interest as the position of each becomes increasingly uncertain. The argument here is familiar and does not require repeating. But it is worth noting that this consideration expands the domain of the feasible and provides us with an independent 'reason for rules'. And this reason for rules, because it involves increasing the likelihood of agreement over the objects of choice, is specifically *contractarian* in character.

2 The Hobbesian Element in Constitutional Political Economy

An obvious benchmark case in discussing 'the reason for rules' and the choice among alternative possible sets of rules, is the case of 'no rules'. If it is to be shown that agents actually *choose* constraints to be subject to, over and above those that are imposed by forces of nature, it needs to be shown why choosing such "facultative" constraints[6] is rational. In analysing the 'no rules' scenario, Buchanan and the CPE tradition have looked primarily to Hobbes. Buchanan is no libertarian anarchist. He takes the necessity of government seriously; and indeed, the construction of government out of the rational agreement of ordinary agents is seen to be a paradigmatic instance of constitutional exchange.[7] The attractions of Hobbes in this connection are self-evident. Hobbes is, on most readings, as close to a *homo economicus* political theorist as it was probably possible to come in the seventeenth century. Hobbes, like Hume, nurtured the ambition—largely shared by modern economics—of developing a theory of human behaviour on all fours with physical science, and with all the rigour and elegance of Euclidean geometry. And Hobbes, exactly like Buchanan, sought to derive an account of the legitimacy of government —and of the exercise of coercive power by government—from an argument based in consent of the governed. But Hobbes, unlike Buchanan, does not develop that argument directly. In the Hobbesian conception, it is an error to impose a requirement that government be directly constrained by popular will. For Hobbes, although citizens plausibly choose to be subject to governmental rule, they cannot choose to be subject to specific laws and policies of their own choosing. The contractarian element in Hobbes plays itself out at some remove from day-to-day political process. And this fact makes the Hobbesian account very different from the standard contractarian one. Indeed, it is not at all clear that Hobbes represents a particularly congenial starting point for the CPE tradition. If one takes Hobbesian anarchy as a point of departure seriously, the subsequent analysis may have to develop along lines rather different from those familiar in CPE discourse.

[6]See "The Relevance of Pareto Optimality" p. 211.

[7]As argued, most notably, in Buchanan's *The Limits of Liberty.*

Let us elaborate briefly. In one early version of the contractarian tradition, associated with Wicksell (1896) and Lindahl (1919) from which much of Buchanan's contractarianism derives, the laws and policies to which people are to be subject are construed as "chosen" directly by the citizenry. Politics is thought of as a direct analogue to n-person exchange. But as Buchanan and Tullock (1962) emphasise, this can only be strictly true if the collective decisions about laws and policies are subject to unanimous consent. And unanimous consent is self-defeating in the sense that a decision rule for collective decisions of unanimity would itself be unanimously rejected. Citizens would foresee that under any system of political decision-making in which each individual citizen holds a veto, virtually nothing would ever get done. Citizens would, so Buchanan and Tullock argue, opt for a less-restrictive decision rule. Under a 66 % majority requirement, for example, citizens might reasonably expect that they would be members of the decisive majority two-thirds of the time; and hence that the benefits enjoyed when in the majority would exceed the expected losses to be made in the one-third of cases when they are in the minority. Being "coerced" into paying for collectively provided projects from which I derive no benefits becomes the price I am prepared to pay in order to "coerce" others to contribute to the collective projects from which I do benefit.[8]

In this conception of politics, there is no 'principal/agent' issue. A 66 % majority rule is supposed to act as a perfect agent for whatever majority of individual citizens happens to constitute the two-thirds coalition. But at least in representative political systems (which represent the overwhelmingly predominant practice), individuals do not vote for policies issue by issue—or even for rival packages of policies: they vote for agents (individual political candidates or political parties) to whom they delegate substantial powers of policy determination. And those candidates will have interests of their own which need not align in all cases with those of the citizen voters—and in some matters can be expected to diverge systematically.

Hobbes's thought involves carrying that case to the limit. The individual or group to whom the power to govern is delegated is treated as having total discretion to act as he/she/they see fit. Individuals are offered a choice concerning whether they are to be 'in-laws' or 'outlaws' but they are not offered a say in what the 'laws' will be—the decision about the content of laws is exercised exclusively by the governing power.

There is one interesting 'half-way-house' in the CPE tradition. This is the case laid out in Brennan and Buchanan's (1980) analysis of the power to tax. In the model of government deployed in that treatment, governments are taken to be

[8]We have placed "coercion" in inverted commas here because there is a question as to whether what is at stake here is 'coercion' at all. It is after all something that by hypothesis all consent to. Perhaps, it can be thought of as the collective equivalent of being forced to pay for the car that one drives off the car lot, or the food one buys in the supermarket. For, note that the (legal) requirement to pay for such things reflects a collectively defined and enforced property order to which the purchaser may not have explicitly consented and perhaps in some (few) cases may not on balance benefit from.

totally unconstrained by in-period electoral processes. Governments simply do whatever maximises the government's (that is, the political agents') preferences. However, in the Brennan/Buchanan model, citizens do have some residual constraining power: they have access to a possible 'constitutional restriction' on the *extent* of the taxing power. This model is a move in the direction of the Hobbesian approach—in that citizens have negligible influence on what governments do. But it falls well short of the full Hobbesian thought experiment because the Brennan/Buchanan "Leviathan" is subject to a broad fiscal constraint that specifies the extent of the 'power to tax'. Hobbes' Leviathan is restricted to constraints (if any) that it is somehow profitable (and feasible) for a rational Leviathan to impose on itself. If the set of constraints that fit this requirement is non-empty, that set will constitute the "Hobbesian constitution". And it is, we hope, perfectly clear that any such set is likely to fall well short of what is envisaged to emerge from the 'constitutional agreement' envisaged in Buchanan and Tullock (1962).

Put another way, constitutional political economy has plausibly constructed a 'reason for rules'. But it has operated as if rules of the kinds over which choice is presumed to be exercised are actually feasible. That is, only if we can take it that we *can* collectively impose on ourselves rules of a kind that we can reasonably trust each to abide by, we can reasonably enquire as to which set of rules is best. If there are no rules that agents will abide by, then although we may all in principle benefit from the imposition of such rules, there is little point in pursuing analysis because rules of this kind are simply infeasible. The whole exercise becomes one in futility. If we are to take the 'no rules' benchmark not just as an imaginative construct but as a plausible social equilibrium—as both Hobbes and Buchanan do—then we have to explain how it is exactly that escape from that equilibrium is possible. How is it possible that 'rules' do the work they are purported to do?

Perhaps one might respond to this challenge with the observation that as far as one can see, people do seem to obey rules in enough numbers to make the assumption of rule-following behaviour plausible, and so the question of the feasibility of rules is itself a rather specious one. But if this is so, then one does not seem to have any account of why government is needed for the enforcement of the chosen rules; or indeed why the whole exercise of justifying the existence of government as an instance of 'the reason of rules' is not self-contradictory. Of course, this issue is a problem for Hobbes, no less than for CPE. And it is by no means clear what exactly Hobbes' own solution to it is; or whether that solution is coherent. However, it is of considerable interest to see how Hobbes tackles this puzzle; and what a plausible Hobbesian solution to it would imply about how the 'constitutional perspective' ought to be applied to specific issues of institutional design.

This is the aim of this paper. We begin with an attempt to construct a version of the Hobbesian argument that seems to us to be at least coherent. We then turn to a consideration of what this would mean for the 'constitutional perspective' and for a genuinely contractarian approach if the Hobbesian Constitution were the relevant one.

It perhaps goes without saying that the particular Hobbesian construction of the state of nature and the possible routes out of it is quite different from other scholars' construction of the 'no rules' case. If the contractarian position took John Locke (1689) or Jean-Jacques Rousseau (1762) as a point of departure, doubtless the challenges of reconciliation would be rather different. But as a matter of fact, it is Hobbes who figures in the CPE literature as defining the benchmark 'no-rules' case; and so it is Hobbes's logic with which CPE has to grapple.

3 The Hobbesian Story

A disclaimer is in order at the outset in this exercise. We are not historians of thought in anything except an utterly amateur sense. We do not claim that our reflections represent "what Hobbes really meant"; we do not see ourselves to be providing an interpretation of Hobbes in that sense. There are a number of extremely interesting (and highly varied) such interpretations. Our approach is to take what we see as various 'fixed points' in the Hobbesian account of emergence from the state of nature; and try to fit a minimalist coherent structure around those fixed points. It is in that sense that we tend to refer to the "Hobbesian logic" rather than to Hobbes *simpliciter*. Arguably, what we are engaged with here might be best described as *Buchanan's* Hobbes—Hobbes, as he figures in framing the broad constitutional contractarian story. For our main concern is with whether the Hobbesian logic as deployed in the CPE tradition, as we read it, is consistent with a standard contractarian story in terms of its implications for constitutional analysis.

There are good reasons for not getting committed to Hobbesian exegesis—quite apart from our own limitations. For example, there seem to be two different treatments of the central issues—one in *Leviathan* and one in *De Cive* (both published in 1651). Moreover, there are multiple "contracts" in play in these discussions, with different terminologies and apparently different concepts in the two treatments.

For example, in *Leviathan* we find two separate contracts:

(i) *the pactum unionis*, through which individuals abandon "moral anarchy" and move to "moral order", in Buchanan's terms, and
(ii) the *pactum subiectionis* through which all but the Leviathan are subdued to rules.

The *pactum unionis* involves individuals agreeing to treat each other as morally equal and to refrain from using the force to solve interpersonal conflicts. But the status of this particular 'contract' is unclear. In particular, while it is clear what individuals forego with the *pactum unionis*, it is much less clear what they actually get. One might for example think that morally equal individuals would feel equally bound by the contract they have agreed on. If that were so, however, the consequence would be that the state of nature is made up of self-constraining individuals.

Depending on the content of such self-restraining rules, life in the state of nature might be rather limited but hardly as nasty brutish or short as the familiar picture of Hobbesian anarchy is taken to be. Indeed, one might think that within a society of moral equals, *all* individuals, none of them excluded, would be equally entitled to take decisions. In this sense, the *pactum unionis* appears to be designed for some kind of *direct or pure democracy.* But in Hobbes's treatment, it seems to be deployed as simply the first step to *direct or pure absolutism.* It seems as if, for Hobbes, the *pactum unionis* was not to be viewed as a contract at all, but merely a sort of preliminary compromise between parties. From Hobbes' viewpoint, then, it is the *pactum subiectionis* that is the real contract with which individuals shift from the state of nature to the civil state. But this contract seems on its face to violate the moral equality stipulation of *pactum unionis* because by construction those who exercise the authority of the civil state lie outside its authority.

De Cive finesses the problems created by the dual contract, but the "patch", which Hobbes uses to do this, seems to involve a radical alteration of the whole theoretical framework. In effect, *De Cive* cancels the *pactum unionis*: there is only the *pactum de imperio*, and the precise relation of this latter contract to the categories used in Leviathan remains unclear.

Accordingly, we shall take a stick figure version of Hobbes—consisting of what we take to be the central features of the account (or what the standard interpretation seems to be in CPE circles)—and construct from that a basic Hobbesian logic.

The relevant fixed points are as follows:

(a) the state of nature is nasty, brutish and short;
(b) the state of nature is [at least] a short-run equilibrium;
(c) exit from the state of nature is feasible only by virtue of establishing a sovereign;
(d) the state under the sovereign is an equilibrium;
(e) the state under the sovereign is better for all than the state of nature.

Note that any theoretical account that can satisfy the requirements of that story for a set of rational agents has several challenges. First, we have to show that the state of nature is a possible equilibrium in some meaningful sense. Second, we have to show that the state with sovereign is an equilibrium in that same sense. Third, we have to show that this latter equilibrium is the *only* equilibrium other than the state of nature. Fourth, we need to show that the sovereign equilibrium is better for all agents than the state of nature equilibrium. And finally, we have to show that the route from the one equilibrium to the other is accessible to agreement among the parties of some plausible kind.

The Buchanan strategy in this setting is to conceive of the state of nature as an n-person prisoners' dilemma. The 'equilibrium' outcome in this game is one in which each individual wages wars of pre-emptive aggression on all others—that is, the strategy of 'universal defection/non-cooperation' is indeed an equilibrium strategy for all rational players. We can also imagine that there might exist an agent who, if appointed to the role of enforcer of cooperative behaviour [or of agreements

by individual agents to engage in cooperative behaviour], both *could* and *would* enforce that behaviour.

But both the 'could' and the 'would' here are problematic. 'Could' first. How is it feasible for the sovereign to enforce the behaviour in question? What alchemy is wrought in the process of making that individual 'sovereign' that gives him the power to enforce that which he did not possess in the state of nature? Put another way, when the sovereign attempts to enforce his will, what makes me obey? And if the answer involves an appeal to the force of a promise I have made, why could I not make a promise to behave cooperatively [contingent on others doing so perhaps] directly with other individuals in the state of nature? If the sovereign can indeed enforce, why do we need him? That is one issue.

What of 'would'? What is it that motivates the sovereign to do anything other than exploit the subjects who have (somehow) placed themselves under his thrall? How is it that agents have done anything other than replaced a state in which life is nasty, brutish and short for one that is nasty, slavish and all too long? Clearly, there must be an answer to this question, because otherwise it would not be rational for agents to leap out of the state of nature in any case other than that in which they stand a chance of being king. Or perhaps the latter possibility *is* the answer. That is, the expected return arises from agreeing to a regime in which someone is chosen at random to be king and the benefit on offer is just the benefit from enslaving all others. But if so, what binds the agent to adhere to the rules once the regime is in place? Once I discover that I am not the king, why should I rationally obey the one who is?

Presumably, Hobbes believed he had answers to these questions, although Hobbes scholarship seems somewhat at odds over what those answers were. In any event, we do not see ourselves restricted to those answers to which Hobbes might have had access. We are content to raid contemporary social science for plausible answers and let those scholars with expertise in Hobbes exegesis explain why it is, or is not, the case that these particular answers were those that Hobbes had in mind.

In the state of nature there is no external enforcement of promises. Hobbes makes clear that the only obligations that can be binding are self-binding ones. It is useful therefore to distinguish at the outset between two classes of promises: those that are self-binding; and those that require external enforcement. Denote these as type-A and type-B promises, respectively. Then it must be the case that the move out of the state of nature is in two steps. First, each makes a type-A promise, which is sufficient to establish the sovereign. However, to directly adopt the 'cooperate' strategy in the n-person prisoners' dilemma must involve a type-B promise—one unenforceable in the absence of the sovereign. Otherwise, the state of nature could not be an equilibrium; and the external enforcement powers of Leviathan would not be required. Second, it must be the case that enforcement of cooperative behaviour under the sovereign does not require fulfilment of a type-B promise by any agent. These two requirements seem to constitute constraints on any coherent rendering of the Hobbesian story.

Consider an example of this structure that seems at least prima facie consistent with Hobbes' own account. The characteristic case of a type-B promise is a promise

that would imperil one's life. Agents simply cannot be relied on to keep promises that involve them in risking their lives. A promise that involved the undertaking to risk one's life in the future would be "unnatural" and incredible. So, in particular, a promise to enter directly into a contractually supported common peace would be of this kind. There is simply nothing to ensure that I will, if I so promise, be exempt from attack because my promising does not ensure the promise being made or adhered to by others. Whenever anyone has reason to fear that he is under threat for his life, he will defend himself to the death, earlier promises notwithstanding. But lesser promises—say, to obey the sovereign's will whenever one's life is not at stake—are credible. So in particular, the promise to pay taxes and do 'public duty' of a non-life-threatening kind according to the sovereign's dictates can be taken to be (self-)binding. Now, suppose that external enforcement involves enrolling appropriately many against one, in such a way that, though the one defend himself to the death, the many are under no risk of death. Say that if the enforcers outnumber the non-cooperator by more than three to one, all of the enforcers are utterly safe. Then all can commit credibly to acting as enforcers of the sovereign's will in those numerical circumstances, though none can promise to abide by the sovereign's will unilaterally—or indeed to never resist attack should one be subject to such. Thus, even should the sovereign decide to execute one, for whatever reason, the one will resist; but the resistance will avail him nothing provided the sovereign has called on support in appropriate numbers.

If this reasoning provides us with grounds for thinking that creation of a sovereign is both necessary and sufficient for a leap out of Hobbesian anarchy, it does not provide us with reasons for believing that this leap will be better for all players. It does not then provide us with grounds for believing that the move will satisfy the contractarian test; and relatedly, does not necessarily explain why choosing to make the necessary type-A promise is rational for all agents. In order to provide such grounds, we need to say something about the rational conduct of the sovereign, once established.

Note first of all that the rationality test for ordinary agents only requires that the state under the sovereign be better for the agent *in an expected sense.* It may be expected that the sovereign will act in such a way as to make some agents worse off than they would be in the state of nature; and agents can know that this will be so for some agents. Provided only that the identity of such agents is unknown at the point where the type-A promise is made, then an expected return calculus is all that is required to make the giving of that promise rational. Whether or not an agent is subsequently made worse off does not set any logical bounds on what the afflicted agent can do. Once the sovereign is established, no single agent can unilaterally plunge society back into a state of nature.[9] So for example there is no constraint that the leap out of anarchy must satisfy the Pareto test (i.e. make everyone better off) for the type-A promise to be rational. But we do have to show that agents have

[9]Buchanan seems to suppose the opposite in his "Before Public Choice"; this seems to us to be a departure from strict Hobbesian logic.

reason to expect that they *will* be better off under the sovereign on average. What grounds do they have for that belief?

Let us suppose that 'support for the sovereign' comes in degrees. In particular, within the terms of any feasibly enforceable law, there is a kind of principal–agent problem for the sovereign in ensuring that his subjects do what would actually be best for the sovereign. Suppose, in particular, that the sovereign has read his Smith and his Hayek. The sovereign knows in particular that the wealth of his nation will be maximised under a regime of only modestly modified laissez-faire with well-ordered markets. Of course, to maximise the wealth of the nation in the Smithian sense—that is, to maximise the aggregate income of his subjects—is of no direct interest to the sovereign. It only becomes 'interesting' to the sovereign to the extent that he can appropriate some portion of that wealth. And there will clearly be a trade-off between achieving maximal aggregate wealth and achieving the maximal amount for the sovereign. But, in particular, it seems possible that the best the sovereign will be able to do will be to levy a uniform tax rate on all income and otherwise allow trade and the division of labour to flourish unabated. In other words, the sovereign will rationally implement a regime of 'natural liberty' in Smith's sense, supplemented with a revenue-maximising fiscal regime. And that revenue-maximising fiscal regime will interfere minimally with the functioning of the economic order because that is the way to maximise fiscal returns. Of course, the sovereign could attempt to manage the activities of his subjects in the manner of a slave–owner managing a community of slaves—that is, specifying the tasks of each and enforcing punishments for any failure to reach specified targets. But the sovereign's familiarity with Hayek reminds him that such 'managed economies' work rather poorly, and that the extensive division of labour characteristic of truly 'wealthy' economies can emerge only when each individual agent is free to respond to market signals and is provided with relevant incentives to do so. It will be useful to have a term for this regime. Let us call it the 'modified system of natural liberty'—MSNL. The characteristic features of MSNL are as follows: each agent has effective property rights in the fruits of her own labour, apart from the share which the sovereign appropriates in taxes; and that there are all the institutional facilities of the free market, including civil protection for all fiscally relevant rights, rules for the free exchange of rights and/or the fruits from them, and enforcement of contracts concerning such exchanges. The 'rights' of all parties against other private parties will be enforced up to net revenue-maximising levels. And given the residual interest that the sovereign has in the economic flourishing of each and every citizen and the economic flourishing of the whole, it seems plausible that the sovereign will have considerable incentive to enforce the rights of citizens against violations by other citizens with considerable vigour.

More perhaps needs to be said about the precise properties of the revenue-maximising tax regime; but if we think of it as involving a simple uniform-rate income tax, set at an appropriately high level [well short of 100 %], this approximation will do for the purposes of the argument. Effectively what the rational sovereign is doing is to devise a structure of private rewards for citizens such that the share left with them is sufficient to induce them to produce for him the

maximum revenue of which they are capable. This is an idea that is familiar from the recent economic literature on rational dictatorships—though of course, for real-world dictatorships, the threat of rebellion creates constraints of which Leviathan in the strict Hobbesian model is exempt by construction. If the sovereign is genuinely secure—if, that is, the type-A promises that agents made in the state of nature are truly binding on citizens—then the sovereign will have no reason to depart from his MSNL regime. That is the regime that it is rational for the sovereign to pursue—or rational, at least, for that sovereign who is motivated by the desire to maximise his economic power, the splendour of his court and his own private fortune. If citizens know all this, then they have reason to trust that their lives will go much better for them on average than in the state of nature. A simple reliance on the sovereign's prudence will give them adequate reassurance of a better future under the sovereign than under a 'no-rules' regime.

However, it is to be emphasised that the sovereign here is not bound by any contract. The sovereign provides no undertakings to citizens in this picture and is not bound to provide any undertakings. The only contractual element is the set of type-A promises made by citizens in extricating themselves from the state of nature to abide by the will of the sovereign. The sovereign himself remains in the state of nature with respect to any undertakings made, namely, that they must be 'self-enforcing'. And so far, no undertakings of any sort by the sovereign have entered the picture. The idea that the sovereign enters into any contract with citizens in the establishment of the Leviathan order has no place so far in the logic of the argument. So, to this point, we have no Hobbesian Constitution: simply predictably rational action on the part of the sovereign, once established.

4 The Hobbesian Contract

A Hobbesian Constitution, if one is to exist, must depend on the sovereign's having reasons to bind himself in certain ways. That is, it must be the case that strictly rational action at every point does not involve things going as well for the sovereign overall as they might if the sovereign were to resile from [the possibility of] rational action in certain contexts. We know that Hobbes thought that this situation could arise. Citizens in the state of nature will do things [rationally] that they would do better to foreswear; citizens *can* bind themselves [if only by type-A promises] and it is rational for them to do so. Since the sovereign remains in the state of nature in the sense that there is no external enforcement, the sovereign has access to type-A promises, which are credibly binding, if he should wish to avail himself of them. Why would he?

Note that the regime of the revenue-maximising tax depends on a measure of trust on the part of individual citizens. For the proper functioning of the MSNL, at least some agents will have to accumulate capital. And they will, to some extent under the uniform income tax regime, reveal information about their income-earning capacities that it would be logically possible for Leviathan to

exploit. That is, Leviathan might be tempted to levy higher tax rates on those who reveal an exceptional income-earning capacity and/or to confiscate the accumulated capital of those whose accumulations reach significant magnitude. Unless the citizens can be sure that Leviathan will not give into such temptations, the possibility will moderate the income-earning/accumulation behaviour of citizens in a way that is income-reducing for Leviathan overall.

Consider for example the single-shot trust game in this setting. Here, citizen [C] gets to choose first whether to accumulate, thereby rendering her liable to expropriation in the next period. If she chooses not to accumulate, she and Leviathan both receive a return of zero. In the second period, Leviathan [L] gets to choose between simply applying the income tax on the return—in which case both C and L receive a pay-off of 2—or to apply a confiscatory wealth tax in which case L gets 4 and C loses 2. Thus, the pay-off structure is depicted in Fig. 1—which is the standard 'trust' game.

The characteristic feature is that there exists an equilibrium in this game—namely the no accumulation outcome—which is Pareto-dominated by another technically feasible outcome—the 'income tax' outcome. However, the income tax outcome is, in this one-shot game, ruled out by Leviathan's rationality. If L chooses the higher pay-off open to him at any choice node, L will choose to expropriate, leaving C worse off than if she had not accumulated. Ergo, C will not rationally accumulate and both C and L will forgo mutually beneficial accumulation activities.

The hypothesis that some measure of accumulation will occur in the MSNL regime, as in the account sketched above, depends on an implicit assumption that the logic in the one-off game will not apply—that L will rationally refrain from appropriation because other C's in future periods would not accumulate. For the indefinitely iterated game, self-restraint may be L's best strategy. But for this to be the case—for L to exercise self-restraint indefinitely—requires assumptions about L's perceived security of tenure and discount rate—assumptions that might be implausible under certain contingencies. And if the likelihood of these contingencies arising is high enough [given the pay-offs in Fig. 1, if the probability of expropriation is greater than 50 %], then C will rationally *not* accumulate. So it will pay L to pre-commit to the strict income tax regime, if such pre-commitment is feasible.

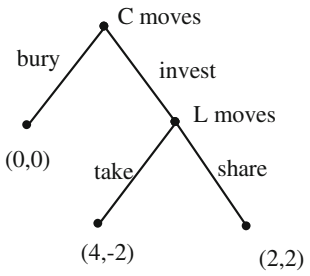


Fig. 1 The standard trust game

One thing we know *is* feasible, that is for L to make promises of type-A. For there is nothing which is available to any citizen in the state of nature that is not equally available to L and if citizens have the wherewithal to escape from the state of nature through credible promises of the type-A, then L must have access to such self-binding as well. One way to think of this issue is to postulate that promises are binding if the temptation to break them is not too great. So, the game indicated in Fig. 1 becomes the augmented game in Fig. 2, in which L has the option of making a promise; and it is common knowledge that breaking a promise costs L something —an amount s in Fig. 2. In Fig. 2, we represent the reward to L from expropriation as E, while retaining all other values as specified in Fig. 1. We can on this basis analyse the Leviathan interaction with citizens and the citizens' emergence from the state of nature by reference to the same general analytic structure and the same two parameters. The first parameter is the value of s, which measures the force of promises made; and the second parameter is E, the value of the temptation at stake in failing to fulfil promises made. Clearly, here if $E > (s + 2)$, then promises will not be adhered to. Type-B promises are those such that this condition is satisfied. This includes specifically promises that imperil one's life. Equally clearly, if $s > 0$, then a promise has some force and promises formally made, even without external enforcement, increase the likelihood that they will be kept across a range of possible contingencies.

Now, L may have access to a number of manoeuvres that serve to increase the likelihood that any constitutional promises made by L will be adhered to. It may be possible for L to ensure that any violation of such a promise is more public and known to be more public, so that the cost in terms of aggregate citizen response to any act of expropriation, for example, is likely to be magnified. There are, however, limits to what the Hobbesian sovereign can do by way of institutional redesign. The Hobbesian sovereign will never rationally cede ultimate power. There can be no real division of powers in the Hobbesian world—because the sovereign can always ensure his will by threat of killing the subject in question. There can be no real ceding of powers back to the people and no contingent contracts between sovereign and people. All there can be ultimately are unilateral actions on the part of the

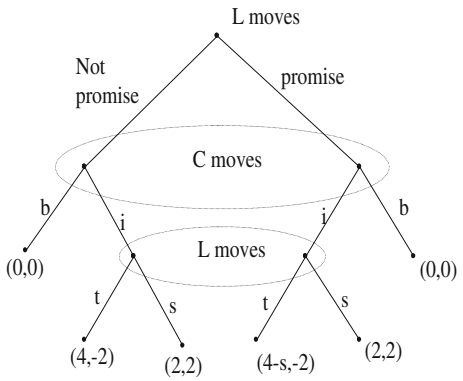


Fig. 2 The augmented trust game

sovereign all of which must be rational, within the limits imposed by access to type-A promises. [The fact that we have modelled those type-A promises as we have allows us to recognise the moral force of the promise, the s-factor, without relinquishing the formal apparatus of rationality. The agent L chooses rationally when he keeps his promise, because the act of promising affects the pay-off structure. There may be things to be said against this formulation—but one thing to be said in its favour is that it exhibits the important property that *behaving* as morality requires is a function of the temptations involved. There is here a downward sloping demand curve for moral behaviour in the sense that the higher the opportunity cost of behaving morally the less moral conduct one will observe.]

The bottom line here is that there will be a Hobbesian Constitution, based on the limited capacity of the sovereign to bind himself through promises unilaterally made. The sovereign cannot, however, be taken to use his 'external' enforcement powers against himself; so the Constitution so formed has real limits from a conventional contractarian viewpoint. There can be no division or separation of powers. Power is one and lies exclusively and unalienably with the sovereign. We might therefore prefer to think of the Hobbesian Constitution as a kind of 'covenant', much like the old and new covenants that God is understood to make with his people in the Christian account—unilateral uncontingent outpourings of 'grace and favour'. In the Hobbesian case, however, the grace and favour have a necessarily instrumental value to the sovereign. They enable the sovereign to enjoy a level of personal well-being that exceeds that to which he would otherwise have access.

If the Hobbesian Constitution involves no separation/division of power, what can it involve? First, for all citizens whose lifetime net contribution to the common wealth [i.e. aggregate taxable capacity] is positive, promises of protection of person and property against action from other citizens. Second, threats of swift and terrible death to any who would threaten the sovereign's authority or person. Third, undertakings by the sovereign to exercise a measure of restraint in the exercise of his power—and specifically, to refrain from violence against citizens and their property, beyond the enforcement of the general citizen obligation to assist the sovereign in the pursuit of the sovereign's ends and the citizen obligation to pay the tax on income imposed by the sovereign. From the sovereign's viewpoint, all of these provisions are of the "kill not the goose that lays the golden egg" variety; and in that sense, in general keeping with the object of having the sovereign's course go as well as possible overall. But the constitution is not externally enforceable. So its force has to lie solely on the 'moral' weight of promises made, and of the sovereign's design of institutions that would make public any failure to keep promises. It seems clear to us that the idea that promises have some [but limited] weight in the state of nature is an essential feature of the Hobbesian scheme. The limits, however, are also notable. In particular, there can be no institutional arrangement that does not leave the sovereign better off; or that requires the sovereign to be subject to any measure of external enforcement (That these are not the same thing is what the trust predicament shows.).

In the Hobbesian world, the *quis custodiet ipsos custodes* challenge bites quite deeply. There can be no arrangement that violates the interests of the *custodes*; and even the advantages to be obtained from self-binding are limited to those accessible under the moral force of ordinary promises. These two constraints obviously impose quite severe limits on the scope and nature of the Hobbesian Constitution. But the constraints do not imply that the idea of a Hobbesian Constitution is a contradiction in terms.

Perhaps the spirit of the foregoing can be captured in terms of a simple diagram. On the vertical axis is the long-term annual consumption of the sovereign; and on the horizontal axis is the annual consumption of the typical citizen. The line NL indicates the notional locus of returns to both sovereign and citizen as the move out of the state of nature proceeds. It can be thought of as showing the locus of various feasible outcomes as the value of the parameter s [the moral force of type-A promises] increases from zero. On our reading of the Hobbesian logic, when s is zero, no escape from the state of nature is feasible. As s increases, two things happen. First, the promises of citizens in pledging support to the sovereign become more robust. And, second, the capacity of the sovereign to make mutually beneficial credible promises in relation to the treatment of citizens becomes more extensive (Fig. 3).

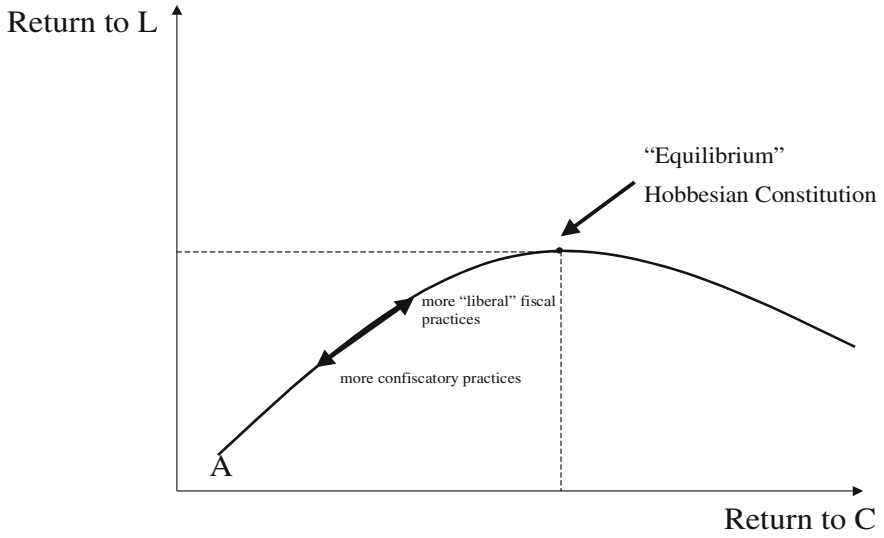


Fig. 3 The limits of sovereign and citizen common interests

5 The Hobbesian/Contractarian Contrast

The argument of this paper can be brought to a suitable conclusion by focussing on the central divergence between Hobbesian and standard contractarian approaches. The standard contractarian approach begins with a depiction of the Hobbesian state of nature as an n-person prisoners' dilemma, and focuses on the potential that all citizens would have for a much improved life under a regime of collectively enforced rules. It poses the question as to what set of rules would be ideal from the citizens' collective viewpoint and seeks to answer that question by imagining what rules would command unanimous approval, after appropriate bargaining/exchange/compromise behind an idealised 'veil of ignorance'. The Hobbesian approach, as we have depicted it here, insists on the imposition of three logical constraints in this entire exercise. First, a set of enforceable rules must be *feasible*. That is, it must be possible to establish a sovereign with the requisite powers. Hobbes takes it that this feasibility test can be met—and it is our conjecture that this is so only if there is some kind of promise/contract that can be made in the state of nature that will prove binding over some range. Second, it must be *necessary* to establish a sovereign. That is, whatever scope for agreements in the state of nature there is, such scope cannot be sufficient to allow directly binding collective contracts—since then the state of nature would not and could not be a genuine equilibrium. Third, it must be the case that the citizens have reason to believe that their lives will go better under a sovereign not subject to external enforcement than in the state of nature itself. We have sketched out an argument as to why this expectation is reasonable, based on the sovereign's capacity to gain some share of the benefits from the escape from the state of nature. Further we have argued that the same limited capacity of promise-making/keeping that enables the citizens to agree to establish an effective sovereign also enables that sovereign to commit to a limited Constitution—a 'covenant' with his people which though not subject to external enforcement nevertheless has some force, and indeed precisely the same force as required to allow the escape from the state of nature in the first place.

Suppose such a Hobbesian regime is in place. It can hardly fail to attract the attention of the contractarian that the sovereign enjoys considerable rents—rents that could in principle be redistributed to the citizenry in a variety of ways. In this respect, the contractarian is no different from the 'pretender' who lusts after the power and wealth that the sovereign possesses. But the contractarian is different in this respect—that if the gains were to be redistributed in certain ways to the citizenry, under an alternative set of institutional arrangements, there would be an increased 'pie' for redistribution. The revenue-maximising income tax regime may be better from the citizens' viewpoint than many alternatives, but it is not the best. The wealth of the nation would be yet further increased under a less-modified system of natural liberty. In this sense, however, the Hobbes-informed observer would see the contractarian as making the same mistake as the 'pretender'—the mistake of failing to recognise the relevance of the feasibility constraints—the mistake of making the good the enemy of the [feasible] best. And as Hobbes sees it,

the cost of that failure is extreme. It threatens a return to the state of nature—a state that hovers whenever citizens forget the moral force of their original promise to support the sovereign up to the limits of the cost of their own lives. The challenge that Hobbes poses for the contractarian is not to show that more liberal regimes than the Hobbesian are *desirable*, but to show how they are feasible—at least within a frame that recognises the state of nature as a relevant point of departure for constitutional analysis.

Put simply, whereas the contractarian assumes that all the gains from the exit from the state of nature are available for appropriation by the citizenry, the Hobbesian recognises that only a limited share of these gains is available. Whereas for the contractarian, the test for desirability is the conceptual agreement of all citizens as equals behind the veil of ignorance, the Hobbesian test must always include the prevailing sovereign as party and include as feasible moves only those changes that are consistent with the sovereign's interest. There are, of course, critical contractarian elements in the Hobbesian account; but those elements operate at the level of having any rules at all, rather than in settling directly the question of what the unconstrained 'ideal' set of rules might be. As we have noted, it might well be possible to offer a contractarian account of the 'reason of rules' from a point of departure other than the Hobbesian one. But once one has chosen the Hobbesian state of nature as the benchmark case of 'no rules', one seems committed to the Hobbesian logic 'all the way down'; and that commitment gives rise to an account of feasible constitutionalism that seems on its face quite different from the emerging CPE orthodoxy in certain crucial respects. Our aim in this paper has been to indicate something of what those "crucial respects" are.

References

Brennan, G., and J.M. Buchanan. 1980. *The Power to Tax: Analytical Foundations of a Fiscal Constitution*. New York: Cambridge University Press.

Buchanan, J.M. 1969. *(Cost and Choice: An Inquiry in Economic Theory*. Chicago: Markham Publishing Company.

Buchanan, J.M. (1982). The Domain of Subjective economics: Between Predictive Science and Moral Philosophy. In *Method, Process, and Austrian Economics: Essays in Honor of Ludwig von Mises*, ed. M. Kurzner, 7–20. Lexington: D.C. Heath).

Buchanan, J.M. (1991). The Foundations for Normative Individualism. In *The Economics and the Ethics of Constitutional Order*, 221–229. Ann Harbor: University of Michigan Press.

Buchanan, J.M. 1999 "The Relevance of Pareto Optimality"; "Positive Economics, Welfare Economics and Political Economy". In *The Collected Works*, ed. James M. Buchanan. Indianapolis: Liberty Fund.

Buchanan, J.M., Tullock, G. 1962. *The Calculus of Consent: Logical Foundations of Constitutional Democracy*. Ann Arbor University of Michigan Press.

Lindahl, E. (1919). Just Taxation—A Positive Solution. In *Classics in the Theory of Public Finance*, eds. R. Musgrave and A. Peacock (1967). London: Macmillan.

Locke, J. 1689. *Two Treatises of Government*, ed. P. Laslett. Cambridge: Cambridge University Press.
Rousseau, J.J. 1762. *The Social Contract*. London: Penguin Classics.
Wicksell, K. 1896. A New Theory of Just Taxation. In *Classics in the Theory of Public Finance*, eds. R. Musgrave and A. Peacock (1967). London: Macmillan.

Reforming the Fiscal Constitution: Holding Politicians Accountable Through Greater Transparency

James W. Douglas, Ringa Raudla and Robert S. Kravchuk

1 Introduction

In the era of growing deficits and fiscal crises, advancing discussions over the rules of the game that should govern the budget process is more paramount than ever. An important insight of constitutional law and economics is that while we cannot change the nature of politicians, we can change the rules of the game—and through that, influence policy decisions and outcomes (for an overview of the literature, see Raudla 2010). Jürgen Backhaus has contributed significantly to re-invigorating the constitutional—or Wicksellian—approach to public finance, as witnessed most prominently by the *Handbook of Public Finance* (Backhaus and Wagner 2006; see also Backhaus 2002; Backhaus and Wagner 2005a, b).

In this chapter, we draw on the insights of constitutional law and economics and constitutional political economy and make a proposal for a fiscal constitutional reform (the term "constitution" here refers to the rules of the game in general, not only to the provisions that are included in the formal written constitution). We argue that in order to restore relevance to the executive phase of the budget process, it

J.W. Douglas (✉)
Department of Political Science and Public Administration, University of North Carolina at Charlotte, Charlotte, NC, USA
e-mail: jwdougla@uncc.edu

R. Raudla
Ragnar Nurkse School of Innovation and Governance, Tallinn University of Technology, Tallinn, Estonia
e-mail: Ringa.raudla@ttu.ee

R.S. Kravchuk
School of Public and Environmental Affairs, Indiana University, Bloomington, IN, USA
e-mail: kravchuk@indiana.edu

A. Marciano and G.B. Ramello (eds.), *Law and Economics in Europe and the U.S.*,
The European Heritage in Economics and the Social Sciences,
DOI 10.1007/978-3-319-47471-7_5

needs to provide better information to citizens regarding the options available for closing the gap between revenues and expenditures. We outline a proposal whereby the executive would have to make clear his/her priorities by designating spending as either revenue or deficit financed, and submitting a 5-year balanced budget plan within his/her budget recommendation. The legislative body (or bodies) would then be required to enunciate how their preferences differ from those of the executive in their respective budget resolutions. We contend that such a process would be more transparent, better enabling citizens to hold politicians accountable. We illustrate the implementation of the proposal with the case of the United States.

In 1921, Congress established an executive budgeting process with the passage of the Budget and Accounting Act. One of its goals was to create a better system of accountability (Willoughby 1927; Caiden 1987; Joyce 2008). Members of Congress complained that presidents frequently blamed the legislative branch for "extravagant" spending without providing either proof to support claims of inappropriate congressional spending or alternative spending plans. The expectation was that an executive budget would provide citizens with a comprehensive plan regarding the nation's finances and better information about who to hold accountable for poor spending decisions (Select Committee on the Budget 1919: 212, 278, 379). The executive budget is, therefore, supposed to serve as an instrument of presidential leadership and voter accountability. Many observers of federal budgeting, however, believe that it has lost much of its relevance and has little impact on overall totals (Caiden 1987; Pitsvada 1988; Rubin 2007; Meyers and Rubin 2011). Perhaps it is time to consider reforming the executive budget process in order to better identify the policy preferences of the president and both houses of Congress. For democracy to function properly, citizens may need a clearer mechanism of accountability than the current system provides.

In this paper, we argue that in order to restore relevance to the executive budget process, it needs to provide more clear-cut information to lawmakers and citizens regarding the options available for addressing the nation's fiscal issues. We believe that a good way to accomplish this goal is to reform the process so that it forces the president, the House, and the Senate to be more transparent regarding their policy preferences relative to one another. Doing so could advance the executive budget process as a more effective mechanism with which to inform citizens about the budgetary consequences of the policy positions of the major actors. The proposal we offer here is meant to begin a discussion of how to think conceptually about reform possibilities, and is based upon what we believe the literature and theory might suggest. It should, therefore, serve as a useful starting point for debate over how decision-making processes might be altered in order to help democracy function better.

The remainder of the paper is organized as follows. We provide a summary of the relevant literature. We then describe our reform proposal. We conclude by explaining why we think our proposal is an improvement over the current process.

2 Information, Accountability, and Tough Choices

Serious problems can occur in democracies when informational asymmetries exist between elected officials and citizens. The concept of fiscal illusion is particularly useful for understanding these dynamics. The notion of fiscal illusion goes back to the writings of an Italian economist Puviani (1897, 1903) from the late nineteenth–early twentieth century and was "re-discovered" by James Buchanan in the 1960s (1960, 1967). As Da Empoli (2002) emphasizes, Puviani's ideas are clearly applicable to modern democratic settings. Indeed, since the 1960s, there has been an increasing attention to the issues of fiscal illusion, as indicated by numerous studies using fiscal illusion both in theoretical modeling and empirical applications (see, *inter alia*, Pommerehne and Scheider 1978; Wagner 1976; Oates 1991; Misiolek and Elder 1988; Dollery and Worthington 1996). The core of the problem of fiscal illusion is that because of lack of information, citizens develop inaccurate perceptions about the levels of taxes they actually pay, the scope of programs they actually benefit from, what the connections between taxes and benefits truly are, how high the actual budget deficits are, and what the sources of the incurred deficits are.

Drawing on the theory of fiscal illusion can be expected that while politicians have better information about the potential consequences of policy proposals, they have an electoral incentive to limit what they release to the public (Persson et al. 1997), and that citizens' fiscal illusions regarding the true cost of government services encourage politicians to behave opportunistically in order to gain electoral advantage. The greater the informational asymmetry, the more opportunistically politicians behave. And, this opportunistic behavior tends to result in undesirable outcomes such as larger fiscal deficits and mismatches between the true preferences of the electorate and composition of the budget (see Eslava 2011 for a comprehensive summary of this literature).

Citizens, therefore, must be adequately informed if good governance is to be achieved. For accountability to work, voters must have the possibility to make rational choices at the ballot box. Hence, they need to know which policy options are available and whom to hold responsible for bad decisions (Besley 2006; Pande 2011). An emerging consensus in the literature examining procedural aspects of budgeting is that accountability in budgeting can be strongly facilitated by institutional elements that enhance the transparency of both the budget process and the budget document itself (see Poterba and von Hagen 1999; von Hagen 2002; Alt and Lassen 2006; Garrett and Vermeule 2008). The more transparent the budget process and the resulting documents are, the lower are the monitoring costs that voters would have to incur; hence, slack in the principal-agent relations would be reduced. Informed voters diminish politicians' ability to behave opportunistically, and should result in better decisions regarding public finances. Furthermore, greater budget transparency appears to result in both higher voter participation and lower fiscal deficits (Benito and Bastida 2009; Eslava 2011). Citizens have shown a willingness to support difficult tradeoffs (even when doing so is not in their

individual economic interests) when provided with good information about the magnitude of fiscal problems and the options available for dealing with them (Tanaka 2007). For example, budget transparency has been shown to raise voters' tolerance for tax increases (Alt and Lowry 2010). Joyce (2008), however, warns that increasing the volume (and even the quality) of information is not *by itself* adequate to produce better decision making. Increasing the volume of information may make things more transparent, but transparency is hardly useful if it overwhelms voters with information. To be useful, information needs to be digestible and make clear and credible tradeoffs between policy options.

Critics of the executive budget process argue that it is no longer serving as a realistic mechanism for setting national priorities and holding policy makers accountable (Pitsvada 1988; Rubin 2007; Meyers and Rubin 2011). Caiden (1987) complains that the effectiveness of the executive budget has been watered down by a drop in the quality of the information it includes. Meyers (2009) and Rubin (2007) argue that the federal budget process in general has broken down. They point out that political opportunism has resulted in budgetary rules being violated and extreme claims being made by elected officials. Meyers and Rubin (2011) argue further that the polarized political environment has caused members of both political parties to avoid making tough decisions that might anger voters, while blaming each other for a lack of action. In this regard, they contend that the executive budget has not served as a useful tool for pushing the agenda forward in a responsible way, stating that "Only rarely have presidents used the bully pulpit to educate citizens about budget problems and propose solutions." (338). Posner (2011) relents that the recent battles over the debt limit, continuing resolutions, and appropriations bills reveal an unwillingness on the part of Democrats and Republicans to turn away from the politics of playing to their political bases in order to compromise in the national interest. What the public needs, according to Meyers and Rubin (2011), is information that allows citizens to expose opportunistic "candidates who promise something for nothing." (342).

Such a mechanism is especially important given Fowler and Margolis' (2014) finding that American voters are relatively uninformed regarding the major political parties' policy positions, and that providing voters with basic information about these policy positions has a significant impact on vote choice. To understand the level of voters' ignorance concerning the federal budget, one need only examine polls showing majorities of citizens citing debt and deficit reduction as a top priority while also opposing most tax increases and spending cuts (Kohut 2012).

Given the claims against the executive budget process, scholars have proposed several reforms in order to enhance priority setting and accountability. Fisher (1990) suggests that the executive budget by itself is, in fact, an effective tool. He lays the blame for much of the difficulties with the federal budget with the congressional budget process, arguing that the budget resolution is the problem because it weakens the influence of the president's budget. He recommends eliminating the budget resolution so that the president can play his traditional leadership role and accountability can be restored as the framers of the 1921 Act intended. The problem with Fisher's proposal is that the executive budget by itself still creates ambiguity.

The numbers in the document are generated by the president's staff (which immediately opens them up to claims of bias), and they reflect the president's policy preferences—which Meyers and Rubin (2011) point out are not necessarily fiscally prudent, and Farrier (2011) demonstrates do not always show strong leadership. Furthermore, Fischer's proposal requires neither the president to address deficit reduction concretely nor Congress to state its preferences clearly. Members of Congress can, as they often do now, simply attack the president's plan as an unrealistic political document while only offering vague statements regarding solutions and blaming the opposite party for unsatisfactory outcomes. Indeed, Congress has often failed to even pass budget resolutions over the past several years, and their absence has not made the executive budget any more effective.

Pitsvada (1988, 1996) argues that the executive budget should be replaced with a budget recommendation produced by a joint executive-legislative committee in order to set parameters that all can work within as the process starts each year. Similarly motivated, Meyers and Rubin (2011) suggest replacing the concurrent budget resolution with a joint resolution. These proposals rely upon the willingness of elected officials to reach compromise. There is little reason to believe that policy makers will produce different results when negotiating on joint congressional or executive-legislative committees; the recent failure of the Joint Select Committee on Deficit Reduction (e.g., the Supercommittee) and the inability of Congress to pass budget resolutions in recent years should be proof enough of that. These proposals do not force elected officials to make their preferences transparent to citizens. While such bi-partisan target setting would be effective in an environment where both sides are willing to compromise, they are less likely to achieve their desired outcomes in the current climate where too many political actors are adhering to extreme positions to gain electoral advantage. The incentive to make extreme claims would still exist, so voters would continue to suffer from confusing and low quality information with which to hold candidates accountable.

What is needed is a system that provides greater transparency regarding the magnitude of the nation's fiscal problems, realistic options for dealing with them, and the policy preferences of our elected leaders relative to one another. The executive budget process in its current form does little to address the informational asymmetry between elected officials and citizens. The president makes a recommendation based upon his policy preferences about revenues and expenditures for the upcoming fiscal year. The executive budget provides estimates of total revenues and expenditures, as well as the resulting deficit/surplus. The president is not required to furnish information about deficit reduction or expansion. When presidents do, their recommendations frequently come in the form of long-term plans. In cases where deficit reduction is the goal, it is in the president's interest to push the costs of such plans as far into the future as possible. This, of course, opens the president up to criticism for not showing leadership on the difficult issues. It, however, is also in Congress' interest to push tough decisions into the future, opening it up to presidential criticism. Both have an electoral incentive to decry the evils of the deficit, but then avoid being the first to suggest a realistic plan for reducing it because they know all such proposals will impose serious costs on the

electorate and be attacked by the other side. On the other hand, deficits are not necessarily bad. Despite both parties' "concern" over structural deficits and the national debt, very little movement is often made on deficit reduction. This is likely because many elected officials believe that *maintaining* or even *increasing* deficits is in the national interest, at least in the short-term. It is rare, however, to hear politicians make this case, leading citizens to believe that deficits are always bad. In the end, citizens are confused about the possible policy choices, the costs of those choices, and who to hold accountable for lack of action. Citizens simply do not know what politicians value most.

In the literature on budgetary institutions, different types of institutional solutions have been discussed. Alesina and Perotti (1999) distinguish between three types of rules that can help to govern the budget process: (1) numerical targets on the budget (e.g., a balanced budget rule or a deficit target); (2) procedural rules outlining the preparation and adoption of the budget; and (3) rules regarding the transparency of the budget. The proposal we outline below addresses all three aspects simultaneously.

3 A Presidential Balanced Budget Recommendation

We argue that the current informational asymmetry between citizens and elected officials can be mitigated by altering the executive budget process. When conceptualizing this proposal, we draw on the insights of Knut Wicksell's classical work on public finance (Wicksell 1896). Although little known among contemporary public finance scholars, Wicksell has been considered to be the intellectual father and founder of modern public finance by *both* James Buchanan and Richard Musgrave (Hansjürgens 2000). As Wicksell emphasized, the institutional setting and the rules of the game within which decisions on public finances are made significantly influence the content of these choices (see, e.g., Backhaus and Wagner 2004). Among other aspects of the budget process, Wicksell pointed to the importance of a simultaneous consideration of revenues and expenditures: in his opinion, budget proposals should always be accompanied by matching revenue or financing proposals in order to secure that expenditures would not be approved without considering the finance requirements they imply (for an overview and discussion, see, e.g., Johnson 2011; Uhr 1951). According to Wicksell, this would allow the voters to compare the benefits of public expenditures and corresponding tax burden, and hence decide whether to support the budget plans (Hansjürgens 2000). Our proposal takes Wicksell's recommendation one step further by allowing voters to evaluate which expenditures should be tax-financed and which deficit financed.

Drawing on these insights and keeping in mind the current context of US federal budgeting, we recommend that the president's executive budget be required to contain three major components: (1) revenue financed expenditures for the

upcoming fiscal year, (2) deficit financed expenditures for the upcoming fiscal year, and (3) a 5-year balanced budget plan.

The revenue financed component would identify the president's highest priority expenditures, those he/she would recommend be funded if the government were required to balance its budget for the upcoming fiscal year. This element of the president's proposals would signal to voters which government activities were most important to him/her, and illustrate just what it is that projected revenues are potentially paying for. Totals would be subdivided into two basic categories: general fund revenues and earmarked revenues. For the former, the president would have unlimited discretion in specifying which government activities he/she would recommend be financed with general revenues. If the president proposed any general fund tax increases or revenue enhancements, he/she would designate which programs he/she would prefer to finance with the projected new revenues. In this manner, citizens would gain some idea of what they are getting in return for any increase in their tax burden.

The proposal would be more constrained for earmarked revenues. Revenues from these sources would have to be allocated to their legally assigned functions. Surplus revenues, however, would be available to finance other activities. For example, the projected Social Security payroll tax revenues would be designated to pay Social Security benefits first. The president would then specify how he/she would prefer to spend the Social Security surplus. This would better inform citizens about what types of things trust fund surpluses are paying for, perhaps dispelling the notion in some taxpayers' minds (perhaps reaffirming it in others') that trust fund monies are being squandered. If, in contrast, the government needed to redeem bonds from a trust fund in order to cover legally required expenditures, then the president will need to indicate how he/she proposes to pay the cost of the bonds—general fund revenues, borrowing from the public, or new revenues? As with proposals for new general fund revenues, the president would have to specify how he/she would prefer to allocate any recommended increases in earmarked taxes.

Any remaining expenditures included in the president's executive budget would be classified as deficit financed. This component of the budget proposal would represent secondary priorities of the president—those worthy of funding, *but not worth the cost of increasing taxes to pay for them*. The president would be required to explain why the country should run a deficit to cover the cost of these programs rather than raise taxes (e.g., recession, war, stimulus, anticipated growth, etc.). This component of the president's budget would inform voters about which programs are relatively low priorities of the president (presumably those he/she would be most willing to cut) and the reasons why it is necessary to borrow to finance them. The sheer size of current deficits would compel the president to identify a substantial number of government functions to include as deficit financed, thus forcing citizens to confront the reality that the deficit does indeed fund programs that many of them support, and that simply "cutting the waste" is not going to be enough to balance the books. Proposed tax cuts would also be accounted for in this section of the president's budget. If the president recommended tax cuts, he/she would be

compelled to propose specific offsets and/or identify which programs would now be financed via borrowing as a result of the loss in revenue.

The final component we recommend the president include in his/her executive budget is a 5-year balanced budget plan. Here, the president would be required to identify what he/she believed to be the most reasonable path to achieve balance (i.e., eliminate the deficit financed portion of the budget) over a 5-year timespan. The multi-year timespan (the length of which could be debated) is necessary to ensure that realistic options are put forward. There are four avenues to balance, each of which would have to be addressed in the proposal. First, the president should identify savings from sun-setting activities—those expected to terminate during the 5-year period (e.g., ending the combat mission in Afghanistan). Second, the president should identify the programs he/she would finance with anticipated revenue growth (above the baseline). Third, the president should identify which of the deficit financed programs he/she would cut. Finally, the president should identify which taxes he/she would recommend be increased as well as how he/she would allocate the new revenues among the deficit financed programs. Of course, the deficit is a moving target, so projected increases in expenditures over the time period would have to be taken into account. The purpose of requiring the president to engage in this exercise is *not* to actually balance the budget, but instead to provide the public with information regarding both the tough choices that must be made in order to balance the budget and the president's specific preferences for achieving balance. Once again, the size of the current deficit would necessitate the president make difficult tradeoffs in his/her proposal, signaling voters about the types of sacrifices that would need to be made to balance the nation's books. If the president truly desired deficit reduction, then the information provided here would obviously map out the president's preferences for achieving that goal. If, on the other hand, the president thought that deficits were in the national interest going into the future, then this information could be used as evidence for why—showing perhaps that the sacrifices would be too great given the economic conditions facing the country.

The president does not have the constitutional authority to make the budget on his/her own. It is therefore essential to require the House and Senate to signal to voters their preferences regarding each of the three components of the budget discussed above. This could be accomplished in each chamber's budget resolution. If, for example, President Obama designated half of defense spending as deficit financed, Republicans in the House and/or Senate might disagree and label all defense spending as revenue financed in their respective budget resolutions. They would, however, have to specify which programs from the president's budget they would redesignate as deficit financed, thus establishing their priorities vis-à-vis the president. By the same token, if President Obama included a tax increase in his 5 year balanced budget plan, House Republicans might counter by replacing it with proposals for cuts to specific programs. Congressional Budget Office (CBO) estimates could be used to ensure that all parties are following the same assumptions. This could be accomplished by requiring the president to respond to CBO's adjustments to his/her initial projections. Conversely, Office of Management

and Budget (OMB) numbers could be used by all parties. While this may invoke complaints that OMB's numbers are politically motivated (as often happens during budget negotiations), research has shown that the OMB and CBO estimates are generally very close to one another (Krause and Douglas 2005, 2006), and what is most important for our reform proposal is that the president and both chambers of Congress make their preferences clear based upon the same assumptions in order to reduce confusion in citizens' minds regarding the differences across those actors' policy preferences.

The House and Senate would have an incentive to actually pass their respective budget resolutions in order to highlight the uniqueness of their policy preferences; something their supporters would likely want demonstrated, especially if they had not been designated as high priorities in the president's budget. Doing nothing would be branded as acceptance of the president's plan as the best course of action for the country, and would show a lack of leadership. Alternatively, to deal with the potential problem of a chamber not passing its budget resolution, CBO could be instructed to establish a default position for any chamber failing to comply with the requirements, where revenue and deficit financing is assigned proportionally across all programs for the upcoming fiscal year (after accounting for earmarked revenues), and balance is achieved via across the board spending cuts and tax increases over the 5-year period. The leadership in the chamber would be forced to defend why its party is unable to improve upon either the president's proposal or the default position. While a chamber still might choose to shirk its responsibility to inform the voters of its priorities; at the very least, the default position would educate voters about how the deficit is being "shared" across programs as well as the costs associated with "sharing" the solution.

Thus, the new executive budget process would provide voters with several vital pieces of information.

- *Accountability.* The president and each house of Congress would be required to indicate which programs and tax policies are most important to them. This would clarify policy preferences, better signaling to voters who to support at the polls.
- *Scale of the Deficit Problem.* Citizens would be shown that the problem of deficit reduction is structural and would require difficult choices.
- *Options.* Citizens would be presented with specific options regarding how to balance the budget, putting them in a better position to assess potential tradeoffs. Citizens would also be shown that a possible "good" option might be *not* reducing the deficit.
- *Simplicity.* Relying on either CBO or OMB estimates would eliminate confusion regarding differences across budget proposals—everyone will be working from the same assumptions. Additionally, CBO could produce a comparative analysis of the proposals, making the major policy differences between the three actors transparent.

4 Conclusion

Politicians behave opportunistically. This is made easier by uninformed voters (Eslava 2011) and complex information (Joyce 2008). Political rhetoric from both parties and branches of government give voters only a vague notion of the nation's fiscal problems and how best to solve them. Elected officials have a political incentive to take extreme positions and avoid compromise (Posner and Sommerfeld 2012). As a result, citizens do not have the information necessary to make decisions about which policies are in the country's best interest. The literature shows that this lack of transparency results in negative fiscal outcomes such as larger deficits (Eslava 2011).

Our proposal to reform the executive budget process is intended to provide citizens with better information and force more meaningful debate and compromise between the political principles. Previous reform proposals assumed that rules can be put in place that will encourage politicians to compromise for the greater good, but those proposed rules would not by themselves induce politicians to make tough decisions. Our proposal, therefore, puts the onus on the voters to force compromise because we assume that such compromise is most likely to be achieved through an informed electorate holding politicians accountable. A properly informed electorate can change the political landscape, inducing elected officials to react (i.e., Fowler and Margolis 2014; Eslava 2011; Alt and Lowry 2010; Tanaka 2007). We argue that this can be accomplished by compelling the president, the House, and the Senate to be more transparent regarding their policy preferences. The current budget process fails to make the preferences of these political principals' and the potential tradeoffs clear and digestible to voters. We believe that our proposal would provide clearer information to voters. The revenue financed component would highlight the programs most favored by the president and show what the government can afford to purchase with the current revenue structure and proposed tax increases. The deficit financed component would identify programs that are of secondary importance to the president, show what citizens are getting when the government borrows money, and justify the need to borrow rather than tax or cut. The 5-year balance budget plan would provide realistic options for the tough choices that have to be made to balance the books—if that is indeed a preferable outcome; and the responses of the House and Senate would make clear the differences between the priorities of the major political principles.

Achieving a balanced budget is a worthy goal. It is a desirable outcome in the long-term, but not the objective of our proposal. Rather, we seek to improve the public's knowledge of both the realities of the deficit and the political parties' primary fiscal policy goals. Balance does not actually have to be achieved. As stated earlier, there may in fact be good reasons why balance is undesirable in the short-term (Posner and Sommerfeld 2012). Knowing how politicians would achieve balance, however, would send a strong signal to voters about what the major players value most—one of the primary goals of the executive budget process as it was originally drafted (Select Committee on the Budget 1919: 212, 278, 379).

We seek to make the differences between the main actors (president, House, Senate) transparent to citizens, by moving beyond vague statements such as "spending is out of control" and "the rich should pay their fair share." *If spending is out of control*, what exactly should be cut? *If taxes need to be raised*, who exactly should pay, and how much? *If borrowing is preferable*, then why and what will it finance?

Ultimately, we do expect a better informed electorate, rather than rules that try to force politicians to compromise, to be more effective at moving the government more toward the black. Research shows that transparency causes citizens to vote their preferences more accurately (Fowler and Margolis 2014), increases voters' willingness to accept tough choices (Tanaka 2007; Alt and Lowry 2010), and actually results in smaller deficits (Benito and Bastida 2009; Eslava 2011). Politicians have an electoral incentive to be vague about their preferences and avoid compromise. Transparency improves accountability, serving as a check on such opportunistic behaviors. In its current form, the federal budget process fails to provide simple and clear transparency. The information provided is overwhelming and confusing (Joyce 2008), and American voters are left uninformed about the political parties' policy positions (Fowler and Margolis 2014). For democracy to function effectively, this must change. We believe that our proposal to reform the executive budget process is a step in the right direction. It draws on the suggestions and lessons from research and theory. We understand that our proposal does not capture all of the complexities of the federal budget process, but it is our hope that in the spirit of previous reform proposals made by other scholars of public budgeting and finance (i.e., Fisher 1990; Pitsvada 1996; Meyers and Rubin 2011), it leads to a larger discussion of how best to improve budgeting for the U.S. government.

References

Alesina, Alberto F., and Roberto Perotti. 1999. Budget deficits and budget institutions. In *Fiscal institutions and fiscal performance* (pp. 13–36). Chicago, IL: University of Chicago Press.

Alt, James E., and David D. Lassen. 2006. Fiscal transparency, political parties and debt in OECD countries. *European Economic Review* 50(6): 1403–1439.

Alt, James E., and Robert C. Lowry. 2010. Transparency and accountability: Empirical results for US states. *Journal of Theoretical Politics* 22(4): 379–406.

Backhaus, Jürgen. G. 2002. Old or new public finance? A plea for the tried and true. *Public Finance Review* 30(6): 612–645.

Backhaus, Jürgen G., and Richard E. Wagner. 2004. Chapter 1: Society, state, and public finance: Setting the analytical stage. In *Handbook of public finance*, ed. Jürgen G. Backhaus, and Richard E. Wagner, 1–14. Boston: Kluwer Academic Publishers.

Backhaus, Jürgen G., and Richar E. Wagner. 2005a. Continental public finance: Mapping and recovering a tradition. *Journal of Public Finance and Public Choice* 23(1/2): 43–67.

Backhaus, Jürgen G., and Richard E. Wagner. 2005b. From continental public finance to public choice: Mapping continuity. *History of Political Economy* 37(1): 314–322.

Backhaus, Jürgen G., and Richard E. Wagner, eds. 2006. *Handbook of public finance*. New York, NY: Springer Science & Business Media.

Benito, Bernardino, and Francisco Bastida. 2009. Budget transparency, fiscal performance, and political turnout: An international approach. *Public Administration Review* 69(3): 403–417.
Besley, Timothy. 2006. *Principled agents? The political economy of good government*. Oxford, NY: Oxford University Press.
Caiden, Naomi. 1987. Paradox, ambiguity, and enigma: The strange case of the executive budget and the United States Constitution. *Public Administration Review* 47(1): 84–92.
Da Empoli, Domenico. 2002. The theory of fiscal illusion in a constitutional perspective. *Public Finance Review* 30(5): 377–384.
Dollery, Brian E., and Andrew C. Worthington. 1996. The empirical analysis of fiscal illusion. *Journal of Economics Surveys* 10(3): 261–297.
Eslava, Marcela. 2011. The political economy of fiscal deficits: A survey. *Journal of Economic Surveys* 25(4): 645–673.
Farrier, Jasmine. 2011. *The Law*: Barack Obama and Budget Deficits: Signs of a Neo-Whig Presidency. *Presidential Studies Quarterly* 41(4): 618–634.
Fisher, Louis. 1990. Federal budget doldrums: The Vacuum in presidential leadership. *Public Administration Review* 50(6): 693–700.
Fowler, Anthony, and Michele Margolis. 2014. The political consequences of uninformed voters. *Electoral Studies* 34: 100–110.
Garrett, Elizabeth, and Adrian Vermeule. 2008. Transparency in the U.S. Budget Process. In *Fiscal challenges: An interdisciplinary approach to budget policy*, eds. E. Garrett, E. A, Graddy and H. E. Jackson. Cambridge: Cambridge University Press.
Hansjürgens, Bernd. 2000. The influence of Knut Wicksell on Richard Musgrave and James Buchanan. *Public Choice* 103: 95–116.
Johnson, Marianne. 2011. Wicksell and the Scandinavian and public choice traditions. *International Journal of Social Economics* 38(7): 584–594.
Joyce, Philip G. 2008. Does more (or even better) information lead to better budgeting? A new perspective. *Journal of Policy Analysis & Management* 27(4): 945–960.
Kohut, Andrew. 2012. Debt and deficits: A public opinion dilemma. Pew Research Center for the People and the Press. (June 14). http://www.people-press.org/2012/06/14/debt-and-deficit-a-public-opinion-dilemma/. Accessed 15 July 2012.
Krause, George A., and James W. Douglas. 2005. Institutional design versus reputational effects on bureaucratic performance: Evidence from US government macroeconomic and fiscal projections. *Journal of Public Administration Research and Theory* 15(2): 281–306.
Krause, George A., and James W. Douglas. 2006. Does agency competition improve the quality of policy analysis? Evidence from OMB and CBO fiscal projections. *Journal of Policy Analysis and Management* 25(1): 53–74.
Meyers, Roy T. 2009. The "Ball of Confusion" in federal budgeting: a shadow agenda for deliberative reform of the budget process. *Public Administration Review*, 69(2): 211–223.
Meyers, Roy T., and Irene S. Rubin. 2011. The executive budget in the federal government: The first century and beyond. *Public Administration Review* 71(3): 334–344.
Misiolek, Walter S., and Harold W. Elder. 1988. Tax structure and the size of government: An empirical analysis of the fiscal illusion and fiscal stress arguments. *Public Choice* 57(3): 233–245.
Oates, Wallace E. 1991. On the nature and measurement of fiscal illusion: A survey. In *Studies in fiscal federalism*, ed. Wallace E. Oates. Aldershot, UK: Edward Elgar.
Pande, Rohini. 2011. Can informed voters enforce better government? Experiments in low-income democracies. *Annual Review of Economics* 3: 215–237.
Persson, Torsten, Gerard Roland, and Guido Tabellini. 1997. Separation of powers and political accountability. *The Quarterly Journal of Economics* 112(4): 1163–1202.
Pitsvada, Bernard T. 1988. The executive budget—An idea whose time has passed. *Public Budgeting & Finance* 8(1): 85–94.
Pitsvada, Bernard T. 1996. A call for budget reform. *Policy Sciences* 29(3): 213–226.
Pommerehne, Werner W., and Friedrich Schneider. 1978. Fiscal illusion, political institutions, and local public spending. *Kyklos* 31(3): 381–408.

Posner, Paul. 2011. *Federalist* No. 30: What is to be done about the federal budget? *Public Administration Review* 71(1): 53–61.

Posner, Paul, and Matthew Sommerfeld. 2012. The politics of fiscal austerity: Implications for the United States. *Public Budgeting & Finance* 32(3): 32–52.

Poterba, James M., and Jürgen von Hagen. 1999. Introduction. In *Fiscal institutions and fiscal performance*, ed. James M. Poterba, and Jürgen von Hagen. Chicago: University of Chicago Press.

Puviani, Amilcare. 1897. *Teoria della illusione nelle entrante pubbliche*. Perugia.

Puviani, Amilcare. 1903. *Teoria dell'illusione finanzaria*. Palermo.

Raudla, Ringa. 2010. *Constitution, public finance, and transition*. Frankfurt am Main: Peter Lang.

Rubin, Irene. 2007. The great unraveling: Federal budgeting, 1998–2006. *Public Administration Review* 67(4): 608–617.

Select Committee on the Budget of the House or Representatives. 1919. *Hearings on the establishment of a national budget system*. Sixty-Sixth Congress. Washington DC: Government Printing Office.

Tanaka, Susan. 2007. Engaging the public in national budgeting: A non-governmental perspective. *OECD Journal on Budgeting* 7(2): 139–177.

Uhr, Carl G. 1951. Wicksell: A centennial evaluation. *The American Economic Review* 41(5): 829–860.

von Hagen, Jürgen. 2002. Fiscal rules, fiscal institutions, and fiscal performance. *The Economic and Social Review* 33(3): 263–284.

Wagner, Richard E. 1976. Revenue structure, fiscal illusion, and budgetary choice. *Public Choice* 25: 45–61.

Wicksell, Knut. 1896. *Finanztheoretischc Untersuchungen*. Gustav Fischer: Jena.

Willoughby, William Franklin. 1927. *The national budget system*. Baltimore, MD: The Johns Hopkins Press.

The Power of Free as a Catalyst for Political Revolution

Dennis W.K. Khong and P.C. Lim

1 Introduction

The Arab Spring of 2011 was a series of recent events that are still fresh in many people's recollection. Unlike previous political revolutions, the events which starting in Tunisia in late 2010 and continuing to other Arab states in 2011 (Anderson 2011), and other subsequent similar events in Hong Kong (Ng 2013) and Taiwan (Rowen 2015) made heavy use of the Internet, social media websites and mobile apps as a means of communications and mass dissemination of message (Lotan et al. 2011).

As reported by various sources, the widespread use of social media websites and mobile apps such as Facebook and Twitter allowed like-minded people to share their thoughts and vent their anger and frustration to an online crowd. Tools which were originally meant to connect people socially were equally adept at connecting people politically. In either case, how the tools are being used and what the messages they carry are determined ultimately by its users.

This paper makes the case that the free and open source software (FOSS) together with free online communication platforms played an important catalytic role in the rise of this modern form of political revolution. The two aspects of 'free', both in terms of freedom to reuse computer codes and downward pricing pressure towards zero marginal cost, expanded consumers' demand and suppliers' competitiveness in the Internet and mobile marketplace. Coupled with this is the positive network effect which increases the marginal value to each consumer as the number of users or subscribers to a particular network increases. The culmination of

D.W.K. Khong (✉)
Faculty of Law, Multimedia University, Melaka, Malaysia
e-mail: wkkhong@mmu.edu.my

P.C. Lim
Support Center for Gender Equality, Yokohama National University, Yokohama, Japan

A. Marciano and G.B. Ramello (eds.), *Law and Economics in Europe and the U.S.*,
The European Heritage in Economics and the Social Sciences,
DOI 10.1007/978-3-319-47471-7_6

these forces is a set of evolving information technology tools to support the expansion of the voice of the citizenry and the growth of political revolutions.

2 Economics of Information

The theoretical starting point for the analysis of free is Arrow's (1962) paper on the nature of information in relation to inventions. Arrow makes the observation that with advancement in reproduction technology, the cost of reproducing information is very low or zero, and as such, efficiency would require "unlimited distribution of the information without cost". This essentially is Arrow's declaration of the zero marginal cost nature of information.

Hence, information, knowledge and what the subject matters of what we usually term as 'intellectual property' are characterised as public goods, or public goods-like, for they exhibit the twin characteristics of non-rivalrous in consumption and non-excludability (McNutt 1999). Non-rivalrous in consumption means that the quantity and quality of information does not decrease with use. This is a good characteristics because it means that information can be reused indefinitely, and coupled with the zero marginal cost of use, as Arrow earlier indicated, would mean that it is socially efficient to allow information to be used by anyone who wishes to use it.

The second characteristics of information is non-excludability. It means that once information has been disclosed to the public, it is very difficult, if not impossible to retrieve it back, and to prevent non-payers from using it. To the creators of new information, especially those of inventive type, this characteristic is bad news because it would be difficult for them to recoup their initial cost in creating or discovering those information.

Due to the difficulty, if not inability, to extract payment from users of inventive information, as described by the characteristics of non-excludability, the market failure of free riding is said to occur when users use information without paying its creators. The modern solution to this free-riding problem is a form of exclusive right, such as copyright, patent or trade mark, to force potential users to pay through a license agreement, on the pain of an infringement suit. The narrative of this intellectual property right assumes that without these exclusive rights, there will an acute under-supply of new knowledge and information goods, which the world needs to fuel its economic development and prosperity.

3 Free and Open Source Software (FOSS) Movement

Proponents of intellectual property proclaim that without intellectual property rights, creators of informational goods will have to compete with free-riding competitors, and since these competitors do not incur the fixed cost of creations, and therefore

need not recoup any investment or repay any financing debts, free-riding competitors will be able to compete downwards towards their marginal cost, which could be zero or close to zero. Under such competitive environment, selling informational goods at zero marginal cost is not a profitable proposition, because that would entail zero producer surplus and zero profit. So it goes that intellectual property rights such as copyright are needed to grant an exclusive right to the creator to deter free-riding competition and offer a chance to the creator to sell the informational goods at a price higher than marginal cost, and thereby make sufficient profit to cover the initial fixed cost of creation and the cost of doing business.

Ultimately, propertisation through intellectual property rights should be seen as only a second best solution. As Arrow (1962) has rightly pointed out, the fact that information is kept proprietary through intellectual property rights is itself a market failure, for efficiency would dictate that information should be made available at its marginal cost, which theoretically is zero. Therefore, the best solution remains at making the informational goods available freely, and for the creator to earn profit through some other associated ways, such as provision of services or sales of physical products.

The idea of giving away something such as information for free and obtaining profit through sale of a physical good is not a recent invention. Anderson (2009) recounted the story of how Jell-O was first introduced to American households through a door-to-door campaign to distribute a free copy of recipe book teaching housewives on how to use Jell-O in their kitchen.

Giving something away for free as a business model also appeared in the market for computer software in the form of freeware and shareware since the early days of microcomputers in the 1980s. Freeware (Crawford 1985), a concept to be contrasted with the later concept of free software, is a software to be given away for free with no expectation of payment. Sometimes, but not necessarily always, the source code of the freeware is also given. The closest legal concept of freeware is the public domain in copyright.

Shareware expands and modifies from the idea of freeware by requesting for voluntary payment if a user continues to use a piece of freely distributed software. It builds upon an honour system, and may prove to be a low cost but effective business model (Hui et al. 2008). Using a technological means, shareware may also implement product differentiation or versioning by disabling certain features unless users pay for an unlock code.

The concept of license-free software has a long history dating back to the first commercially available computers, the mainframes of the past. In the olden days, mainframe and microcomputer manufacturers see themselves as retailers and makers of computer hardware. Software was just something that is to be given away to make the hardware work. Computer programs on paper tapes were made freely available and circulated among user communities. It was only later that software is seen as a separate business of its own, to be developed by software houses with no direct ties to the hardware makers (see Gates 1976).

Notwithstanding software later acquiring a business life of its own, the idea that software, especially source codes, should be made freely available did not die. Richard Stallman, through his Free Software Foundation campaigned for a concept called free software through their flagship model license termed the GNU General Public License (GPL) so conceived in 1986 (Williams and Stallman 2010, p 127). According to Stallman, a software is said to be 'free software' if it enjoys four freedoms:

(i) The freedom to run the program as you wish, for any purpose,
(ii) The freedom to study how the program works, and change it so it does your computing as you wish. Access to the source code is a precondition for this,
(iii) The freedom to redistribute copies so you can help your neighbour, and
(iv) The freedom to distribute copies of your modified versions to others. By doing this you can give the whole community a chance to benefit from your changes. Access to the source code is also a precondition for this.

Key to Stallman's idea is not just about computer programs without a price, but the accessibility of computer codes for study, modification and reuse. This allows programmers to 'stand on the shoulders of giants', to do more and better. Thus there is no need to reinvent the wheel every time, and programmers, under free software licenses, can reuse tested and functional codes developed earlier by others. To Stallman, the freedom aspect of 'free' is more important than the *gratis* aspect of 'free'.

The use of the term 'free' by Richard Stallman is problematic because the word 'free' has many different meanings. The most common meaning in English for 'free' is 'no price'. When the word 'free' is used to describe an animated entity such as a person or an animal, it has the meaning of 'unconstrained' or freedom. On the other hand, when the word 'free' is used in relation to an inanimate object, we cannot be referring to its freedom aspect, but an unpriced aspect of the object, i.e. *gratis*. Therefore, Stallman's use of the phrase 'free software' to mean freedom is an unnatural use of the word 'free' and was a constant source of confusion.

When one of the major Internet software companies, Netscape, made an announcement in January 1998 to release the source code of their browser program under a form of free license, Raymond (1998) and a few other Internet luminaries decided that a rebranding effort of the term 'free software' is necessary to make the concept friendlier and acceptable to businesses, for they are afraid that their businesses would not be able to profit from software if their software is 'free software'. After tossing around some choices, they settled with 'open source' as the agreed replacement term for 'free software'. Many prominent members of the computing industry in the US and the media have accepted this replacement, although Richard Stallman continues to launch an attack against the usage of the new term (Stallman 2009), on the ground that his 'free software' movement is arguing a different value from that of the 'open source' movement.

Whether free software is the same as open source software is debatable. The Open Source Initiative, which manages the 'Open Source' trademark, publishes a list of ten criteria for the definition of open source software, which is largely similar to the free software definition published by the Free Software Foundation. Nevertheless, there remains some minor differences in the recognition of some software licenses on whether they are free software or open source software, although in recent years, the term 'open source' is more commonly used than the other, and in the academic literature, the use of the term 'open source' in lieu of 'free software' is almost ubiquitous.

4 Open Source Licensing

In simple terms, open source software is software licensed under a license which allows public access to the source codes of a software. These licenses rely on domestic copyright law for recognition and enforcement, and so, they tend to be written in generic enough terms such that they comply with the norms established under international copyright convention, namely the Berne Convention for the Protection of Literary and Artistic Works.

One of the most used pieces of software licensed under an open source license is the Linux kernel. The Linux kernel can be described as the central component of a modern operating system which facilitates the communication between application programs, data and hardware. It was the final missing piece to completion of a free operating system started by Richard Stallman called the GNU Project. The Linux kernel is a piece of software initially released by Linux Torvalds in 1991. It is licensed under the GNU GPL version 2, and therefore is free software under FSF's definition.

One of the benefits of releasing the Linux kernel under GPL is that it is free to be reused by others as long as any modifications are released back to the public under similarly terms, namely GPL version 2. Licenses which require 'licensing back' are known as a viral or copyleft license. The advantage of a viral or copyleft license is that it ensures the pool of source codes licensed under the same license grow over time.

5 Android Operating System

Given that the Linux kernel exhibits maturity and robustness and confers freedom under GPL, it was natural for Google to select the Linux kernel for the development its Android operating system for mobile devices such as smartphones and tablets. This decision saves Google a substantial amount of time, effort and expenses in developing a complete operating system for smartphones to compete with Apple's

iOS. In essence, FOSS licensing offers a low cost means of 'standing on the shoulders of giants'.

Other components of the Android operating system developed by Google are also licensed under the Apache License version 2.0 (APL), which the Free Software Foundation determines is a free software license and is compatible with GNU GPL. It is different from GPL on two main aspects. First, it is not viral in nature, which makes it easier to merge source codes under different free and open source licenses. Second, it requires granting of a patent license from a contributor if the software infringes the contributor's own patent.

The combination of GPL and APL for licensing the software of the Android operating system proves to be a potent combination. It allows rapid deployment and adoption of the Android operating system by smartphone manufacturers. A license to use the Android trademark could be obtained from Google subject to several other conditions. But the fact that the Android operating system itself is licensed under open source licenses means that the cost of obtaining a license is low, if not free. Further the licenses also allow smartphone manufacturers to customise the Android operating system to provide unique user interfaces and support specialised hardware configurations.

Since the research and development cost for smartphone manufacturers have been brought down through freely available stock Android operating system, manufacturers could pass on this cost saving to consumers through cheaper smartphones. Also, the pressure of price competition in Android smartphones ensures that the price of Android smartphones is not prohibitive.

6 Law of Demand Meets Smartphones

Classic Econs 101 classes teach the law of demand as the inverse relationship between price and quantity demands by consumers. Thus, when price goes down, the quantity demanded goes up. The law of demand therefore explains widespread usage of Android smartphones when manufacturers engage in intense price competition. Unlike Apple's iPhones operating system iOS which requires expensive research and development to develop, a large part of the development cost of the Android operating system has already been borned by Google and other open source developers. This allows different manufacturers to engage in price competition in the market of largely a homogenous product, i.e. Android smartphones.

Nevertheless some level of product differentiation is introduced by the Android smartphone manufacturers through variations of components such as screen size, pixel density, processor type, RAM size, storage capacity, camera resolution and battery capacity. Other manufacturers such as Samsung added additional hardware innovation such as an S-pen stylus technology. Another effect of competition based on a largely homogenous Android operating system is the short cycles of new versions of Android smartphones being introduced, with each successive generation offering faster and better features at comparable price to the last. And all these are

due to a phenomenon known as Moore's Law. According to Moore (1965), the Intel co-founder's observation, the number of transistors crammed into a piece of integrated circuit has been doubling on average every 18 months, which serves as an indication of an exponential growth in computing power. The result is technology firms chasing and self-fulfilling Moore's Law and pushing out new and more advanced technology in ever shorter cycles. Prices of Android smartphones are therefore kept low through forces of competition, and phones based on previous generation of technology are priced even lower when newer phones are introduced. For example, the average selling price of Android smartphone has dropped by half from USD441 in 2010 to USD219 in 2015; whereas the average selling price of Apple's iPhones remains the same between USD650 (2013) and USD710 (2011) (Elmer-DeWitt 2016).

The beneficiaries of competition in the market for Android smartphones are price-conscious consumers in developing countries, such as in Asia, Middle East, Africa and South America. Quarterly analyses by DeviceAtlas (2015a, b) support this finding that countries with majority Android users are mainly represented by developing countries with lower GDP per capita; whereas Apple's iOS are favoured mainly by Western developed countries. The lower priced Android smartphones allow less wealthy consumers to have access to the advanced technology, apps and information through a smartphone just as their wealthier counterparts in developed countries do. In other words Android's open source licensing plays a role in getting smartphone technology into the hands of everyone.

7 The Power of Free

Hardware without software is quite useless for smartphones; therefore, smartphones must rely on apps and access to the Internet to be a useful gadget. Therefore, to achieve widespread and ubiquitous use of smartphones, not only must the price of the phones be low enough, apps must be accessible price-wise. Fortunately open source licensing again comes into play in the smartphone apps market.

Almost all Android apps are developed using the Java programming language first release by Sun Microsystems in 1995, although implemented based on an open source implementation of Java that was developed separately from Sun's. In an ongoing intellectual property dispute between Oracle Corporation—the current owner of the official implementation of the Java programming language—and Google regarding the use of Java by Google in Android, a jury trial at the District Court for Northern District of California recently held on 26 May 2016 that the similarities between Google's implementation of the Java's Application Program Interface (API) and Oracle's is exempted from copyright protection under the fair use doctrine. This decision has a tremendous positive impact on the Android ecosystem, as it ensures that Google and all Android apps do not need to pay a licensing fee to Oracle for using Java in their apps. Indeed, with the change of ownership of Java from Sun to Oracle, the reference implementation of the latest

Java programming language on the desktop has shifted from Oracle's to an open source implementation known as OpenJDK (java.net 2016). This has the effect of ensuring that programs written in Java are free from licensing encumbrances, and therefore keeping the cost low.

With the Android operating system being open sourced, the Java programming language license-free, and the free availability of Android development platforms such as the open sourced Eclipse and Android Studio Integrated Development Environment (IDE), the infrastructural entry cost into the Android apps marketplace is fairly low. What is needed for a programmer to develop an Android app is the knowledge of programming using the Java language and the know-how of developing on the Android platform.

By essentially giving the Android operating system away for free, Google hopes to earn its income through other means. The two largest sources of income for Google using the Android operating system are income from advertisements and commissions from the Google Play Store (Petrovan 2016). Since its earliest days, Google monetises its web search facility which it provides for free, by charging advertisers for placing inline text advertisements in search results. Lewis (2010) nicely explains this: "If you are not paying for it, you're not the customer; you're the product being sold". It is the extensive users' profiles and search histories which become the income generating tool for Google in this case.

The second major source of income from Android is the Google Play Store. The Play Store is the default online marketplace for Android apps and digital content such as music, movies, ebooks and digital serials subscriptions, on an Android device. It is a great money-making machine which collects 30 cents commission for every dollar paid through the Play Store. So when billions of dollars are paid for apps and content developed by third-party providers, Google gets a cut of 30 % out of each deal, which amounts to a lot of money.

Given these two sources of income require a large user-based, it only makes sense to allow and acquire as many Android users as possible. And for this reason, the Android operating system is given away for free, for that is the best strategy under this business model.

Google is not the only one who uses this idea of users being the product successfully. Traditional media advertisements on the same principle, for example, television programs, are broadcasted over the air for free. In turn advertisers pay the broadcasters to insert advertisements in the program. Therefore, television viewers are not the customers but the product.

Given that deriving income from advertisements in a free product may be even more lucrative than selling a product, app developers have also gone on to the same bandwagon. Many apps of similar features are given away without charge on the Google Play Store. In this case, the apps will embed an advertisement which gets changed with new content through the Internet.

The choice of programming language to develop Android apps too plays a role. According to the creator of the Android operating system, Andy Rubin, Java was selected as the language for Android because of the easy availability of programmers familiar with the Java language which allowed the Android ecosystem to have

a better chance of growing quickly (Patel 2012). This was proven through by the growth of the size of the Android Play Store. For example, a year after Android was unveiled to the public in 2008, there were about sixteen thousand apps on the Android Market (the previous name for App Store). By the second year, the number has grown to one hundred thousands, and by the third year, almost four hundred thousands. As it stands in 2016, there are more than a million apps in the Play Store.

Facebook is the most recognisable social network website to date. Its service can be access either through a browser or a dedicated mobile app. It is free to use by everyone and does not charge anything to use. Paradoxically, it is valued at USD350 billion. Advertisers pay to advertise on Facebook because of the effectiveness of its targeted advertisements which is made possible through the extensive data-mining of its users' postings and activities on Facebook.

As giving a valuable service away for free and earning through advertisements is such an effective business model, almost all social network websites and apps of any significance adopt the same pricing strategy. This includes Twitter, the 140-character messages broadcasting service; WhatsApp, an instant messaging platform for smartphones; Instagram, a photo and video sharing platform; and YouTube, a video sharing website. Many of these free web and app services proved to be influential in the rise of political revolutions such as the Arab Spring and the democratic protests in Hong Kong and Taiwan.

It can be conceived that if smartphones were more costly, social networking apps and services have to be paid for, Internet is less accessible, and then the impact of information technology on the political revolutions would be felt less. The Arab Spring may even not have happened because information will be restricted to that sanctioned and controlled by the state media such as national broadcasting. Grassroot protest activities may be more difficult to organise, especially when there was no prior organisation planning for such eventualities.

8 Network Effect

Social network and social media platforms such as Facebook, Twitter and WhatsApp are the best modern manifestations of network effect. Direct network effect refers to the phenomenon of an increase in the value to users and would-be users as the number of users to a network, such as a social networking platform, increases (Klemperer 2008). Traditional example of a direct network effect is the telephone. Essentially, a network with only a single telephone unit is of no utility as it cannot be used to phone another person. But the value of the telephone network increases as more telephone units are connected to the network, as users can now have more persons to phone to.

The same principle applies to social network platforms. Indeed, the past has shown that user growth rates for Facebook, LinkedIn and Twitter are not linear, but exponential (Håland 2011). This supports the idea that as the number of users in a social network platform increases, more would-be adopters would join, and this

creates a feedback loop which leads to an exponential growth rate. The accelerated growth due to network effect may lead to early platforms enjoying a first mover advantage given similar features to later platforms. The reason for this can sometimes be attributed to the presence of switching cost and a lock-in effect (David 1985). Once a user has chosen a network where all his friends and acquaintances are on the same network, it would be difficult for him to switch to another network with less friends. Indeed this seems to be also a coordination problem, although Android phones do not restrict the installation of only one social networking app.

The network effect on social network and social media platforms can be both a boon and a bane to early platforms. Since friendship and acquaintances tend to be confined to people of the same generation, users from a later generation may not be so attracted to an existing platform, and may choose to select a different platform closer to the temperament of their generation. A recent survey among young social network platform users found that the adoption of Facebook has declined among this generation of users, whereas use of other social network platforms such as Snapchat and Instagram has exploded (Plummer 2015). Unlike the case of the QWERTY keyboard, there is no persistence in the choice of a social network across different generations.

9 Democratisation of Mass Media

Websites and apps such as Facebook and Twitter are variously characterised as social network and social media. These two terms actually refer to two distinct but related concepts. Social network refers to the facility to link up friends, acquaintances and like-minded groups. For most users, this would mean keeping their privacy setting of their messages at private, so postings are only limited to those whom the users know. On the other hand, the same platforms are also social media, in that they allow individuals to broadcast their views to the public or a selected group of followers.

It is the function of social media which is a game changer in the sphere of mass media. Traditional mass media such as newspapers and the electronic media can be described as communication of few to many. In such a case, the media can be captured and controlled by powerful interest groups such as political entities and wealthy business interests. The messages that these interest groups send out might not be what the minority population or people at the political fringe are interested, or they may be at what Anderson (2006) calls the long tail. Or in less democratic countries, all the major media entities are controlled either directly or indirectly by the ruling government.

Social media changes the balance of power in all these. Social media platforms empower everyone and anyone to be a broadcaster, and as such the model of communication can be characterised as many to many. Thus, social media is a democratising mechanism of the Internet technology. For example, Facebook allows the creation of public pages which allow the owners to post messages

readable by all followers and the public. Twitter may be configured as open to the public and is extensively used by politicians in the US during the recent presidential election primaries to spread their messages to their followers. The fact that tweets may be re-tweeted allows the same messages to be further spread to users on the platform who are not the original followers of those tweets. Furthermore, social media offers an avenue for the politicians to engage in an online discussion with followers and opponents without having to rely on the platforms controlled by traditional media.

More importantly, social media on smartphones provides a different accessibility experience than in the past, for information and messages can be broadcasted instantly and received almost simultaneously by followers. Prior to the rise of social media websites, there had been some Internet services which allow for many-to-many communication, such as mailing lists, newsgroups and websites. Social media on smartphones allows updating of information in real time which played a crucial role in spreading up-to-date information on the ground during the days leading up to the Arab Spring revolutions. Also, in the case of protest movement in Hong Kong and Taiwan, it facilitated coordination in relation to supplies and logistics of protesters who were occupying the area outside the Central Government Offices. When attempts were made to cut the mobile data facilities, protesters turned to another free app called FireChat which uses mesh networking technology to pass messages around from one nearby phone to another without having to rely on traditional Internet connectivity (Smith 2014b).

Another interesting case of using of social media to spur political action was recently observed in Japan. On 15 February 2016, an anonymous poster on the *Hatena Tokumei Diary* (Hatena Anonymous Diary) bemoaned about her inability to get her child registered in a nursery due to a longstanding shortage of spaces in nurseries throughout Japan (Anonymous 2016; Nonomiya and Oda 2016). When Prime Minister Shinzo Abe brushed aside the complaint on the ground that it was an anonymous posting and that its veracity could not be ascertained, a groundswell of support appeared in Facebook and Twitter (Osaki 2016). Online protest eventually turned offline with angry parents protesting in front of the Japanese National Diet. A *change.org* petition also received more than 20,000 signatures and support. As a result, the Prime Minister promised to take action to improve children's nursery facilities by increasing places by 50,000 by 2017.

10 Codes Want to Make You Free

Stewart Brand, an American writer and the editor of the pre-Internet era Whole Earth Catalog, is attributed as the creator of the famous meme "information wants to be free" (Brand 1985). When Brand first coined this phrase, he was using the word 'free' in its financial context, because he was observing the decreasing cost of transmitting and duplicating information digitally. This conforms with what Arrow said two decades earlier (Arrow 1962).

What is more significant is that the word ‘free’ has adopted a secondary meaning, that is, freedom. This is logical given the way ‘information’ in the meme has been phrased as a subject, such as ‘the bird’ in the example of “the bird wants to be free”. To metaphorically speak of information wanting to gain freedom means not only that it is socially desirable for information to be widely accessible without restriction, but also that the use of the information, such as where it is use and how it is use, should be without control. One of the most important uses of information is the transformative use. A transformative use of information is the use of a piece of information to be incorporated into another to make a bigger or better piece of information, in whatever form it may be. Transformative uses of information are prerequisite of progress in society because it saves on duplicative research cost and allows “standing on the shoulders of giants” (Turnbull 1959, p 416). In addition, transformative use preserves freedom of the human civilisation as much of what we do and create is based on borrowing prior knowledge and information of the past. Thus the freedom of information enforces an intergenerational quid pro quo for the greater good of the society.

Computer codes are a recognised form of information, albeit as instructions for machines instead of for the mind. Just as information wanting to be free, since the beginning of the computer revolution there has always been a subculture of computing enthusiasts treating computer codes as free. The Hacker Ethics mandates code sharing among computer programming communities (Levy 1984), and it is the essence of this hacker ethics which has been incorporated into the free and open source licenses.

The numerous benefits of sharing and reusing computer codes should not be understated. The obvious one is cost saving. Open source software allows a developer to build upon an existing software and thus allows saving of time and effort from building from scratch and promotes efficiency. Secondly, open source software also promotes quality codes, as tried and tested codes can be reused instead of writing potentially buggy codes which have to be debugged later. Also, there is a saying in the open source community that “given enough eyeballs, all bugs are shallow” (Raymond 1999, p 30), which suggests that given that open source software can be scrutinised by many developers and bug hunters around the world, no bugs will persist for a long time because someone would soon be able to come up with a fix for it. Thirdly, open source licensing fosters collaborations among programmers and also between users and programmers as users are empowered to log bug reports and suggest features, and are even encouraged to contribute in any way to make the software they use better. Last but not least, since open source software may lower the cost of writing software, developers could then offer the software at a cheaper price or even for free.

The reality of the matter is that open source software has created a culture or a movement to de-commodify software. Several major open source projects have successfully produced software of high quality which matches, if not surpass, the quality of proprietary software. Examples of these include Mozilla Foundation’s Firefox browser and Thunderbird email client, LibreOffice’s office suite, Blender 3-D animation software, the GIMP image manipulation software, the GNU/Linux

operating system, Apache web server, numerous programming languages and tools, and many more. At the same time, it has also open up a culture of providing a slew of free-to-use services on the Internet, such as Facebook, Twitter, WhatsApp, Instagram and others. With these free tools and services, the world has indeed become smaller and closer-knitted. Communications between citizens of a country and among users of various groups on the Internet are cheap and easy. Just like the lowering of Coase's transaction cost (Coase 1960), new ventures and initiatives could be formed where previously the cost of communication was prohibitive.

11 Power of Free Social Media

There are various scales and stages of revolutions, ranging from peaceful demonstration, protests, riots to political revolution. These revolutions, though differing in scales, are orchestrated with an ultimate aim to change the status quo of a political economic, as well as a societal and cultural setting. Over the years, we witnessed a change not only in the roles played by the activists which operated through organisations, but also the tools of communication used in organising such revolutions.

As is evident, social media increasingly plays an important if not decisive role in political change and revolution. Although the scale of some protest might not be large, free social media had successfully helped civilian to protest and forced the government to respond to their requests.

As the incidences in Hong Kong, Taiwan and Japan illustrate, not only are political changes and resolutions catalysed by free social media in the less developed and tumultuous regions, it seems that they can equally work their magic in the more developed and democratic world. In the *Hatena Tokumei Diary* incidence, the almost unpredicted protest was stirred up simply by a strongly worded message by a desperate mother of a young child. Harsh words helped to catch attention and ignite a ferment of support. Had no harsher words being used, there would be no protest taking place. Other disappointed parents would have to quietly bear the inconvenience and disappointment of non-admission. Solidarity was formed with other parents recognising a shared predicament. It could also be argued that social media alone did not instigate a revolution but it allowed deprived groups to unite under a shared view. Neither was it necessary that the message had to have an intention to provide any protest.

Whether government's actions and responses under pressure of protest is genuine or otherwise is debatable. It is possible to take a sinister view and believe that the incumbent government's reaction to the negative social media messages is merely to prevent those messages from going viral and jeopardising the governing party's image and popularity. On the other hand, social media no doubt has become a legitimate platform for the citizenry to put pressure on the government for positive action.

Free Internet tools such as Facebook and the Japanese 2channel have enabled one to contest conventional views delivered through mass media and to criticise national and regional politics at the grassroots level. Social networking operates well in more conservative societies such as Japan where straightforward expression of one's personal views is often constrained by the fear of personal humiliation (Maslow 2011). Furthermore, a survey reported that usage popularity of tweeting functions outweighs that of obtaining news and information (Kawai et al. 2011).

12 Information and Effectiveness of Social Media

In this age of neoliberalism, relations between state and society undergo great transformation. As economic freedom grows and is maximised, a correlated reduction of state intervention is witnessed. Governments lose its domination in controlling information as the society evolves with the rising number of so-called digital citizens. As a result of advanced information technology and innovations, people could easily acquire and are exposed to an abundance of uncontrollable information from various sources other than from the government. With free services on the Internet, the cost of such media services has been lowered to mostly zero.

Relationship between the state and society is very much transformed through political revolutions, both of grand and minor scales. In addition to other conventional tools, free social media is widely used in organising a revolution nowadays. Activists choose social media to organise protests for a couple of reasons. First, it is the low cost of organising whereby physical movement to meetings and so forth is kept to a minimum, that the service is free, and the coverage is wide. Another reason is that it often allows the organisers to remain anonymous and free from governmental harassment.

Nonetheless, activists face the risk of social media betraying the original aims of a protest. In movements which are not sufficiently prepared, they run the risk of messages in social media spreading like wildfires, and thus subverting the leadership of a movement while opening it to a broader membership, or putting its survival in danger, as seen in the Iranian Green Revolution (Papic and Noonan 2011). Many a times, the inability of authenticating the truth of a message could put the social media's reliability into questions. Furthermore, free flow of posted information could lead to biases at times.

Studies have found that social media moderates opinions of users as they are exposed to more points of view (Andrews 2014). However, the diversity of voices produces effects of reducing mass political polarisation as some users refrain from sharing their political opinion on social media, fearing that others may disagree with them or 'unfriend' them over political differences. Surveys however find that for many people in emerging and developing nations, online political dialogue leads to discoveries of shared political leanings of people they know, which is particularly common in sub-Saharan Africa and Latin America (Hampton et al. 2014).

Surprisingly, a recent survey in the US reveals that while we find that social media spurs more information seeking behaviour regarding political or social issue, most of the social media users never discuss politics online or find it frustrated and annoyed when things turn political (Smith 2014a).

Free social media enables protests to take place more easily at a domestic scale. As it is a free tool for reformers to organise political revolution at a very low cost, a movement can depend less on outside funding, and so could create the perception of being a purely indigenous movement and one with wide appeal (Papic and Noonan 2011). However, the success story of political protest aided by social media depends very much on the Internet penetration rates. How successful the revolution movement depends also on how much it had appealed to the middle class and the working class. Even if the social media service is made free and that the populace have the freedom to access, they still need the gadgets and facilities to utilise such tool for organising a political protest. Therefore, low cost of devices such as Android smartphones and cheap or free Internet access are still very crucial in the above equation.

13 Limits of Free Social Media in a Revolution

Free social media has its limits in various ways. First, it is an effective tool in triggering a revolt. However, at stages of negotiation for change that follows, free social media does not facilitate or guarantee effective negotiations. Its episodic nature is usually effective in the initial stage. As we have seen in the Japanese child-in-waiting list case, the government had made a blanket commitment to increase the number of nursery. However, detail problems such as where, what and how they are going to do it remain a great concern of the victims.

Secondly, social media is unlikely to be of much help in the negotiation process. In the Japanese case above, as the blogger was an individual, she is unlikely to be competent in negotiating as a counterpart with the government for change. Thirdly, while social media continues to play essential roles in spreading and sharing information, protesting and condemning governance, other means are needed to support such efforts. Traditionally tools such as organisation and leadership are needed for performing change in the process of revolution. The role of media is also imperative in the modern era. Technology can enhance and facilitate but not replace the fundamentals of organisation (Rahaghi 2012). Without the actual talking and meeting face to face, it is unlikely that communication only through free social media can bring about effective and unified actions in revolutions.

Fourthly, we could study the intended outcome of revolution brought about by social media. In Japan, for instance, protests on major issues such as security, economic and diplomatic usually attract more attention from the government. The voices of the public through actions and visual means could lower the approval rating of the government, which might precipitate the downfall of the incumbent government. However, it is not certain that dissenting voices on social problems,

such as the aforesaid case, rebutted through free social media could lead to the same effect as a change in government. Free social media has the potential to create short-term uprising; it is not certain how it could produce long-term effect for a gradual change of the society.

14 Social Media Usage by Governments

At the present times, freedom of social media is not totally guaranteed by all governments. Governments which traditionally enjoyed the sole dominating power of information, both in attaining and disseminating information to the public, have started to lose control and influence as free Internet usage becomes more widespread. Many governments, however, still monitor, manipulate and practice controls of information. Social media could well serve as intelligence collection tool by the government, to keep track of individuals endangering national security for instance. Governments could also cut off websites and Internet access to prevent and tone down riots, like what happened in China during the riot Xinjian incident in 2009.

While social media is used to mobilise political protests and revolution by the people, governments still has the ability to curb and submerge protests. Nonetheless, in most cases, shutting down Internet itself alone does not reduce or erase the protester's rage. Rather, such acts result in adverse effects, such as exemplifying distrust and disagreement against the government. The *Arab Social Media Report* pointed out that when the authorities attempted to block out information during the Arab Spring uprisings, the authorities ended up spurring people to be more active, decisive and to find ways to be more creative about communicating and organising (Huang 2011).

Finally, foreign policy may also play a role in promoting freedom of using social media. For example, the US has been incorporating into its foreign policy the promotion of the freedom of social media which includes the freedom to access information, the freedom of ordinary citizens to produce their own public media and the freedom of citizens to converse with one another, in line with its promotion of democracy (Shirky 2011). It is hoped that with the push to recognise the importance of free social media by a major superpower, democratic revolutions through the online digital world would be possible.

15 Conclusion

The article highlights and elaborates on the effect of lowering of the cost of software and communication on the Internet and on mobile apps, to the point of being virtually free, to political discourse and political revolutions. It is suggested that the

open source software movement may have played an important but less acknowledged role in the revolutions in recent years.

Free social media could be used as a tool for inciting revolution, to spark changes most effectively at the initial stage. At stages that follow, as social media continues to disseminate information and spread influence and voices of protests, other factors would be needed to support social media, to attain changes. The power of media is one of the important factors; organisation and leadership could dictate negotiating with the governing body so as to bring about transformation of the status quo.

Finally, to take this idea of 'free' to its next logical step, we can assume that whenever the cost of a socially useful good is reduced to zero, new and unexpected uses will arise through the ingenuity of the human race. Free is not necessarily a market failure, for free may be the price to catalyse the next social, economic and political revolution of the human civilisation.

References

Anderson, C. 2006. *The long tail: Why the future of business is selling less of more*. Hyperion.

Anderson, C. 2009. *Free: The future of a radical price*. Hyperion Books.

Anderson, L. 2011. Demystifying the Arab spring. *Foreign Affairs* 90(3): 2–7.

Andrews, N. 2014. Social media can moderate user's politics. *Wall Street Journal*, 21 Oct 2014. http://blogs.wsj.com/washwire/2014/10/21/study-social-media-can-moderate-users-politics/.

Anonymous. 2016. *Hoikuen ochita nihon shine!!!* (I couldn't get nursery. Die, Japan!!!). In *Hatena Anonymous Diary*. http://anond.hatelabo.jp/20160215171759.

Arrow, K. 1962. Economic welfare and the allocation of resources for invention. In *The rate and direction of inventive activity: Economic and social factors*, pp 609–626. USA: Princeton University Press.

Brand, S. 1985. Transcript of Hackers' Conference. *Whole Earth Review* 49.

Coase, R.H. 1960. The problem of social cost. *Journal of Law and Economics* 3: 1–44.

Crawford, W. 1985. Common sense and free software: Or, if it's free, can it be any good? *Library Hi Tech* 3: 47–56.

David, P.A. 1985. Clio and the Economics of QWERTY. *American Economic Review* 75: 332–337.

DeviceAtlas. 2015a. *DeviceAtlas mobile web traffic report Q1 2015*. https://deviceatlas.com/sites/deviceatlas.com/files/pdf/DeviceAtlas%20-%20Mobile%20Traffioc%20Report%20Q1%202015.pdf.

DeviceAtlas. 2015b. *DeviceAtlas mobile web traffic report Q3 2015*. https://deviceatlas.com/sites/deviceatlas.com/files/pdf/DeviceAtlas%20-%20Mobile%20Traffic%20Report%20Q3%202015.pdf.

Elmer-DeWitt, P. 2016. The Apple iOS and Android price gap just gets wider. In *Fortune*. http://fortune.com/2016/02/15/apple-android-asps/.

Gates, B. 1976. An open letter to hobbyists. *Computer Notes* 2(1): 2.

Håland, L. 2011. *Google defies exponential growth, hits 50 million users in 88 days*. https://plus.google.com/112418301618963883780/posts/Jd7fH4C6kHp.

Hampton, K., L. Rainie, W. Lu, M. Dwyer, I. Shin, and K. Purcell. 2014. Social media and the 'spiral of silence'. Pew Research Center, 26 Aug 2014. http://www.pewinternet.org/2014/08/26/social-media-and-the-spiral-of-silence/.

Huang, C. 2011. Facebook and Twitter key to Arab Spring uprisings: Report. *The National*, 6 Jun 2011. http://www.thenational.ae/news/uae-news/facebook-and-twitter-key-to-arab-spring-uprisings-report.

Hui, W., B. Yoo, and K.Y. Tam. 2008. Economics of shareware: How do uncertainty and piracy affect shareware quality and brand premium? *Decision Support Systems* 44: 580–594.

Java.net. 2016. *Java platform, Standard Edition 8 reference implementations*. https://jdk8.java.net/java-se-8-ri/.

Kawai, D., M. Amano, M. Ogasahara, Y. Hashimoto, H. Komuro, S. Oono, and Y. Horikawa. 2011. SNS依存とSNS利用実態とその影響 (Actual use of social networking service and its effects with social networking service addiction). In: *Proceedings of the 26th Annual Conference of Japan Association for Social Informatics*, pp 265–270. https://www.jstage.jst.go.jp/article/jasi/26/0/26_0_265/_pdf.

Klemperer, P. (2008) *Network goods (theory). New Palgrave Dictionary of Economics*, 2nd edition, 5:915–917.

Levy, S. 1984. *Hackers: Heroes of the computer revolution*. USA: Anchor Press/Doubleday.

Lewis, A. 2010. User-driven discontent: Comment. In: *MetaFilter*. http://www.metafilter.com/95152/Userdriven-discontent#32560467.

Lotan, G., E. Graeff, M. Ananny, D. Gaffney, I. Pearce, and D. Boyd. 2011. The revolutions were Tweeted: Information flows during the 2011 Tunisian and Egyptian revolutions. *International Journal of Communication* 5: 1375–1405.

Maslow, S. 2011. Nationalism 2.0 in Japan (http://www.2ch.net/). *Asian Politics & Policy* 3(2): 307–310.

McNutt, P. 1999. Public goods and club goods. *Encyclopedia of Law and Economics* 1: 927–951.

Moore, G.E. 1965. Cramming more components onto integrated circuits. *Electronics* 38: 114–117.

Ng, A. 2013. Fighting inequality in Hong Kong: Lessons learned from Occupy Hong Kong. *Radical Criminology* (2). http://journal.radicalcriminology.org/index.php/rc/issue/view/2.

Nonomiya, L., and S. Oda. 2016. How an anonymous blog went viral and forced Japan's Abe to respond. BloombergTech, 15 Mar 2016. http://www.bloomberg.com/news/articles/2016-03-14/blogger-s-die-japan-tirade-twists-abe-s-arm-on-child-care.

Osaki, T. 2016. Viral anonymous blogger rails against Abe failure to alleviate chronic day care shortage. *Japan Times,* 17 Feb 2016. http://www.japantimes.co.jp/news/2016/02/17/national/social-issues/viral-anonymous-blogger-rails-abe-failure-alleviate-chronic-day-care-shortage/#.V_pieR8tphE.

Papic, M., and S. Noonan. 2011. Social media as a tool for protest. *Security Weekly*, 3 Feb 2011. https://www.stratfor.com/weekly/20110202-social-media-tool-protest.

Patel, N. 2012. Andy Rubin's testimony from the Oracle v. Google trial (round two!). In: *The Verge*. http://www.theverge.com/2012/4/24/2967246/live-andy-rubin-testimony-oracle-google-trial.

Petrovan, B. 2016. How does Google make money from Android? In: *Android Authority*. http://www.androidauthority.com/how-does-google-make-money-from-android-669008/.

Plummer, Q. 2015. Facebook losing teen audience. Now it's just old people socializing. In: *Tech Times*. http://www.techtimes.com/articles/30603/20150203/faceook-losing-teen-audience-now-its-just-old-people-socializing.htm.

Rahaghi, J. 2012. New tools, old goals: Comparing the role of technology in the 1979 Iranian Revolution and the 2009 Green Movement. *Journal of Information Policy* 2: 151–182.

Raymond, E.S. 1998. *Goodbye, "free software"; hello, "open source."*. http://www.catb.org/~esr/open-source.html.

Raymond, E.S. 1999. *The cathedral and the bazaar: Musings on Linux and open source by an accidental revolutionary*. Cal: O'Reilly.

Rowen, I. 2015. Inside Taiwan's Sunflower Movement: Twenty-four days in a student-occupied parliament, and the future of the region. *Journal of Asian Studies* 74: 5–21.

Shirky, C. 2011. The political power of social media: Technology, the public sphere, and political change. *Foreign Affairs*, Jan/Feb. https://www.foreignaffairs.com/articles/2010-12-20/political-power-social-media.

Smith, A. 2014a. Pew Research findings on politics and advocacy in the social media era. *Public Affairs Council*, 29 Jul 2014. http://www.pewinternet.org/2014/07/29/politics-and-advocacy-in-the-social-media-era/.

Smith, J.I. 2014b. FireChat: *The app that fueled Hong Kong's Umbrella Revolution. Observer*. http://observer.com/2014/10/firechat-the-app-that-fueled-hong-kongs-umbrella-revolution/.

Stallman, R. 2009. Why "open source" misses the point of free software. *Communications of the ACM* 52: 31–33.

Turnbull, H.W. (ed) 1959. *The correspondence of Isaac Newton: 1661–1675*, vol. 1. UK: Cambridge University Press.

Williams, S., Stallman, R. 2010 *Free as in freedom (2.0): Richard Stallman and the free software revolution*, 2nd edition. Free Software Foundation.

Regulatory Networks, Legal Federalism, and Multi-level Regulatory Systems

Wolfgang Kerber and Julia Wendel

1 Introduction

Networks of regulatory agencies play an increasing role in the complex governance structures of multi-level regulatory systems. Especially interesting are transnational regulatory networks, in which regulatory agencies from different countries are collaborating for solving regulatory problems. One example is the International Competition Network (ICN) as an entirely voluntary and informal network of competition authorities from all over the world. In the EU networks of regulatory agencies of the member states play an important role within the European regulatory system, which in many policy fields encompasses regulations and regulatory agencies both on the EU and the member state level. Levi-Faur (2011) has shown that in 22 from 36 regulatory fields in the EU at least one active regulatory network existed in 2010. Two important examples are BEREC (Body of European Regulators for Electronic Communication) and ECN (European Competition Network) as regulatory networks in the telecommunication sector and in competition law. The European regulatory networks have been the focus of theoretical and empirical studies in the political science literature, both in regard to their roles within the European systems of regulation and in regard to their specific advantages and problems as a new form of governance in multi-level regulatory contexts (e.g., Eberlein and Grande 2005; Coen and Thatcher 2008; Blauberger and Rittberger 2015).

W. Kerber (✉) · J. Wendel
School of Business and Economics, Philipps University of Marburg, Marburg, Germany
e-mail: kerber@wiwi.uni-marburg.de

J. Wendel
e-mail: julia.wendel@wiwi.uni-marburg.de

A. Marciano and G.B. Ramello (eds.), *Law and Economics in Europe and the U.S.*,
The European Heritage in Economics and the Social Sciences,
DOI 10.1007/978-3-319-47471-7_7

In this article we want to analyse regulatory networks in multi-level systems of governance from a law and economics perspective. Based upon the economic theory of legal federalism, which focusses on the optimal vertical allocation of competences in a multi-level legal system (Van den Bergh 2000; Kerber 2008), we want to ask which role networks of regulatory agencies can play in two-level systems of regulation as present in the EU. In contrast to most of the political science literature, which views regulatory networks primarily as a second-best solution in comparison to the optimal centralisation of regulatory powers—at the EU level (e.g., Eberlein and Grande 2005; Blauberger and Rittberger 2015, 369), the economic theory of legal federalism can show that there are often complex tradeoff problems between the benefits and problems of purely centralised or decentralised solutions. Therefore optimal solutions might consist in sophisticated combinations of centralised and decentralised regulatory powers. Our claim in this paper is that regulatory networks might be an institutional innovation that can help to optimise the tradeoffs between the benefits and problems of centralisation and decentralisation. Drawing upon the many insights of the political science literature about regulatory networks we want to show that regulatory networks can fulfil a number of functions which allow for a better combination of the advantages of centralised and decentralised regulatory powers. In that respect this paper can be seen both as a contribution to the law and economics of legal federalism by introducing regulatory networks as an additional intermediate institutional solution between centralisation and decentralisation, and to the political science literature on regulatory networks for analysing them from the perspective of the economic theory of legal federalism. From that perspective we also claim that regulatory networks should not be seen primarily as transitory phenomenon, rather they can also be a valuable part of an optimal two-level system of regulations in the long run.

In Sect. 2 we present a brief overview of the research upon regulatory networks, esp. in the political science literature, from which we derive four different functions that regulatory networks can play in two-level legal systems. In Sect. 3 we analyse the potential benefits of regulatory networks from the economic theory of legal federalism by explaining how these functions can help to optimise the tradeoffs between centralisation and decentralisation. This theoretical analysis will be complemented in Sect. 4 by three case studies about BEREC and ECN (as European regulatory networks) and the ICN (as a global regulatory network). Brief conclusions can be found in Sect. 5.

2 Regulatory Networks: A Brief Review of the Literature

2.1 What Are Regulatory Networks?

Networks can be seen as institutions that consist of a number of entities, as e.g. firms, agencies, or organisations, and which facilitate coordination and cooperation

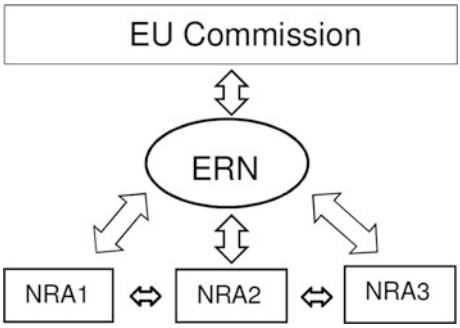


Fig. 1 European regulatory networks as part of European two-level regulatory systems; *ERN* European Regulatory Networks, *NRA* National Regulatory Authority

between these entities. From an institutional economics perspective, networks are a specific group of hybrid organisational structures between hierarchy and market (Powell 1990). In this article we are focusing on the group of transnational networks of regulatory agencies, i.e. that the entities of the network are regulatory agencies from different states.[1] Especially within the specific governance context of the EU, a large number of transnational networks of regulatory agencies of the Member States have emerged and play important roles within the European regulatory system. Maggetti and Gilardi (2011, 1) have defined European regulatory networks as "transnational groups that allow national regulatory authorities to formalise, structure and coordinate their interactions pertaining to the governance of a number of important domains, such as banking, securities, insurance, electricity, gas, telecommunications, broadcasting and competition". However, regulatory networks should not be viewed only as dealing with horizontal coordination problems between regulatory agencies, because they can also play an important role in regard to vertical coordination problems in a multi-level regulatory system (see Fig. 1).

Empirically, forms, characteristics, and functions of European regulatory networks differ widely (Levi-Faur 2011). They show a broad variety in regard to their emergence, (voluntariness of) membership, (informal or formal) organisational structures, independence, competences, and stability. Some regulatory networks (as the IRG, i.e. Independent Regulators Group, in the telecommunication sector) were initiated only by national regulatory agencies and run entirely independent from the EU Commission, whereas others (as BEREC) were initiated from the EU level. Typically, the decision rules of networks are flexible and informal, and membership is voluntary. However, European regulatory networks are increasingly getting institutionalised and formalised (Levi-Faur 2011, 813). In a number of regulatory networks, also the EU Commission itself is a member with certain rights or has at least an observer status. Some regulatory networks have own regulatory powers, whereas others do not play any legally defined role in the European regulatory

[1]Not included are regulatory networks, which also encompass private organisations as firms or NGOs (e.g., as part of private regulation).

regimes. Also the organisational structures of regulatory networks can be very different and change over time as well as one regulatory networks can be replaced by another. In contrast to agencies, networks typically have no own administrative or independent financial capacities (Levi-Faur 2011, 813); however, some of them rely on the budget of a separate office, financed by contributions of the Commission and the Member States (Batura 2012, 6 for the example of BEREC).

2.2 *Regulatory Networks as Governance Instruments in the Political Science Literature*

European regulatory networks have been an important research topic in the political science literature.[2] The theoretical and empirical studies in political science about European regulatory networks can be viewed as part of the broad stream of studies on the specific problems and forms of governance within the complex and unique institutional and political multi-level structure of the EU (see, e.g. Marks et al. 1996; Héritier 2003; Börzel 2010). Due to the difficulties of making political decisions on the EU level, traditional forms of governance as regulation through legislation were partly replaced or complemented by other, new forms of governance. Examples are the use of soft law and the "Open Method of Coordination" (OMC). The basic idea of the OMC was to trigger a process of convergence of policies in fields, where the competences were still largely at the member state level, by establishing a process of identifying best practices and making policy recommendations to the member states (Borrás and Jacobsson 2004; Arrowsmith et al. 2004; Zeitlin 2005; Kerber and Eckardt 2007). Important characteristics of these new modes of governance were, on one hand, their more informal and voluntary ("soft") nature (in contrast of traditional governing through "hard law"), and, on the other hand, their flexible (and also experimental) use in a complex multi-level governance context.

Many of the political science contributions to European regulatory networks start with the assumption of a "regulatory gap", i.e. that the EU is not capable of implementing the necessary effective and harmonised regulatory regimes, because too many regulatory powers still exist at the member state level (e.g., Eberlein and Newman 2008, 26). Since the solution of centralisation of regulatory powers has often not been politically feasible, one of the most important claims of this literature is that the European regulatory networks should be seen as a soft instrument for achieving a stronger harmonisation of the regulatory activities of the member states (Eberlein and Grande 2005; Blauberger and Rittberger 2015). Therefore the EU Commission is often identified as initiator of such regulatory networks (Coen and Thatcher 2008),

[2]See, e.g., Dehousse (1997), Eberlein and Grande (2005), Coen and Thatcher (2008), Eberlein and Newman (2008), Maggetti and Gilardi (2011), Levi-Faur (2011), and Blauberger and Rittberger (2015).

sometimes as a direct response to its failure of establishing a European agency due to the resistance of the member states (Simpson 2011 for the case of BEREC). The political science literature also deals with other research questions, as, e.g. the evolution of these European regulatory networks. Studies in this field have shown that regulatory networks are getting more formalised over time (Saz Carranza and Longo 2012), and based upon a broad empirical investigation, Levi-Faur (2011) claims that European regulatory networks have been increasingly replaced by (European) agencies or are themselves subject to a process of agencification ("agencified networks"). In an econometric study Maggetti (2014) showed that the participation of a national agency in a regulatory network has positive effects on the increase of its regulatory powers but not necessarily on its (organisational) growth. This touches interesting questions about the effects of being a member of regulatory networks in regard to strengthening the independence of national regulatory agencies (Danielsen and Yesilkagut 2014). However, it also raises serious concerns about their accountability vis-à-vis the national governments and parliaments (Lavrijssen and Hancher 2008).[3]

An important part of the political science research is focused on the analysis of the role as well as the advantages and problems of regulatory networks (network governance). In the following, we will structure this discussion by distinguishing four different functions that regulatory networks might fulfil:

1. **Rule-making**: Most of the studies on European regulatory networks emphasise their role in developing and improving regulations. Since the regulatory networks themselves usually have no direct powers for rule-making, their role lies primarily in influencing the rule-making process at the EU level, especially by providing expert advice (Coen and Thatcher 2008). The regulatory networks can have superior regulatory expertise, because they can draw on the knowledge and experience of the national regulators (especially due to their closeness to national markets). However, regulatory networks can also influence the rule-making at the national level, as they participate in the amendment of regulatory frameworks. Therefore, networks can provide the national regulatory agencies with an increased regulatory rule-making capacity and stronger political role vis-a-vis the formal national rule-making institutions as the government and parliament (Danielsen and Yesilkagut 2014, 354; Maggetti 2014, 481).
2. **Best practices and policy learning**: Since the national regulatory agencies usually have developed different regulatory practices in regard to their domestic markets, regulatory networks can also fulfil an important role as a forum for

[3]This is also connected to the view that European regulatory networks are in an area of conflict through a double delegation problem (principal agent theory), resulting from delegating authority from the national level to (1) the EU Commission, and (2) to independent national regulatory agencies (Coen and Thatcher 2008, 51–54).

mutual policy learning between the national regulatory agencies. In European regulatory networks this has often triggered processes of benchmarking and identifying "best practices" (as within the directly related "Open Method of Coordination"). This might lead to regulatory guidance for the national regulators, e.g. in form of norms, standards, and guidelines (Maggetti and Gilardi 2011). In that respect, regulatory networks can fulfil also an important role as channels for policy diffusion (Gilardi 2012) and as breeding ground for regulatory experimentation (Sabel and Zeitlin 2008). In contrast to the OMC, which is generally not seen as very successful (Arrowsmith et al. 2004), best practices and benchmarking procedures in regulatory networks are viewed as a successful soft governance instrument that has contributed significantly to more harmonised rules on the domestic level (Eberlein and Grande 2005; Eberlein and Newman 2008; Maggetti and Gilardi 2011).

3. **Effective enforcement**: Political scientists have emphasised the importance of European regulatory networks in regard to the effective and consistent implementation and enforcement of European regulations (Eberlein and Newman 2008, 26; Blauberger and Rittberger 2015). European regulatory networks seem to be particularly important in policy areas, where the EU has strong regulatory competencies but its operational capacities are weak (Blauberger and Rittberger 2015, 370). Through monitoring regulatory networks can help to close the "regulatory gap" in regard to an effective and equal enforcement of common rules throughout the EU. In that respect, the EU Commission has also been characterised as an "orchestrator", which uses the soft governance instrument regulatory networks as "intermediaries" for influencing the national regulatory agencies (Blauberger and Rittberger 2015). However, regulatory networks also facilitate a more effective enforcement by providing well-established channels for information exchange, communication, and coordination between the national regulatory agencies, building mutual trust between the participants of the network, and allowing for more flexible and effective regulatory solutions (Eberlein and Newman 2008; Radaelli 2008, 243; Maggetti 2014).
4. **Conflict resolution**: Since disputes might exist both horizontally between the national regulatory agencies (e.g., due to geographical spillovers) and vertically between the EU Commission and national regulators, it also can be asked whether regulatory networks also contribute to the resolution of such conflicts. In the political science literature this has been addressed only indirectly, e.g. by emphasising mutual trust, communication, and coordination through regulatory networks (Eberlein and Grande 2005; Sandström and Carlsson 2008), all of which facilitate conflict resolution.

3 Vertical Allocation of Regulatory Powers: The Role of Regulatory Networks

What might be the possible role of regulatory networks from the perspective of the economic theory of legal federalism? In contrast to most of the political science literature on regulatory networks,[4] the economic theory of legal federalism would not assume that the centralisation of regulatory powers or the harmonisation of regulations at the European level is the first-best solution. What the best allocation of regulatory powers in a two-level system of regulations is and how such a system should be designed institutionally, can only be determined after an analysis from a legal federalism perspective. Using also the insights of the political science literature, we want to ask in this section whether regulatory networks can also be a part of optimal institutional solutions in two-level regulatory systems which try to combine the advantages of centralised and decentralised regulatory powers.

Based upon the extensive literature on the economic theory of federalism, which has developed a set of economic criteria about the optimal allocation of competences for public goods and taxation in a federal system (overview: Oates 1999), the economic theory of legal federalism asks more specifically for the optimal allocation of regulatory powers in a federal multi-level system of legal rules and regulations.[5] There are a number of economic arguments which favour more centralisation of regulatory powers and harmonisation of regulations, whereas others emphasise the advantages of decentralisation and regulatory diversity (in much more detail: Kerber 2008, 75–85). For example, the consideration of information, transaction, and regulation costs usually leads to arguments for harmonisation. Different national regulations might also lead to negative welfare effects due to non-tariff barriers to trade or distortion of competition (leading to problems for the Internal Market in the EU). However, if either preferences and policy objectives connected to a regulatory problem or the extent of market failure problems differ between member states, then decentralised regulatory powers might allow for more efficient regulatory solutions than a uniform European regulation. Regulatory powers on the member states level might also allow for the development of better regulations, if the national regulatory agencies hold better knowledge about specific regulatory problems (decentralised knowledge) and/or lead to more regulatory innovation and mutual learning through more regulatory experimentation (laboratory federalism). An additional crucial question refers to the possible advantages and problems of regulatory competition, which can emerge in two-level systems with at least some degree of decentralised regulatory powers.

[4]Some political authors present regulatory networks also as a panacea, see e.g., Slaughter (2005).

[5]See for the relevant economic criteria and the analysis of regulatory competition, e.g., Sun and Pelkmans (1995), Garcimartín (1999), Van den Bergh (2000), Heine and Kerber (2002), Pelkmans (2006, 36–52), Van den Bergh and Camesasca (2006, 406–417), Kerber (2008), and the contributions in Esty and Geradin (2001) and Marciano and Josselin (2002, 2003). For the links to the subsidiarity principle see Kirchner (1997) and Backhaus (1998).

What general conclusions can be drawn from the theory of legal federalism about optimally structured two-level systems of regulations as in the EU (Kerber 2008, 85–87)? A first insight is that the optimal result depends crucially on the type of regulation and the specific regulatory problem. For different regulatory problems, the advantages and disadvantages of centralised or decentralised solutions usually differ widely. This will lead to different optimal vertical allocations of regulatory powers. A second insight is that nearly always significant tradeoff problems between the advantages and problems of centralised and decentralised solutions can be expected. Both insights are also true for the question whether (certain types of) regulatory competition can be expected to yield on balance more beneficial or more problematic (or even disastrous) effects. An important consequence is that most often neither a purely centralised or decentralised solution is optimal, rather the most promising solutions might be found in intermediate solutions, which try to combine advantages of centralisation and decentralisation in a sophisticated way (for contract law: Kerber and Grundmann 2006). This can be achieved in different ways: One possibility is to split the regulatory powers in a regulatory field between the EU level and the member states, i.e. that about some aspects the regulatory power is at the EU level whereas in regard to others it is at the member state level. Another possibility is the separation of rule-making and their enforcement: A centralisation and harmonisation of a regulation might be combined with a decentralised enforcement of these (European) rules, e.g. by national regulatory agencies. In the following, we want to show why also regulatory networks might be a specific type of such an intermediate solution that helps to optimise the tradeoffs between centralisation and decentralisation.

Rule-making: From the perspective of legal federalism there are advantages and problems, if rule-making is allocated either at the EU level or at the member state level. Networks of regulatory agencies in a two-level system of regulation can help to mitigate the problems of solutions, which are either primarily centralised or decentralised. If it is deemed as necessary to have a strong European regulation with a tendency to harmonised rules, then regulatory networks of national agencies might be a very helpful institution for getting access to decentralised knowledge and experiences of the national regulatory agencies about regulatory practices and the specific problems and market conditions in different member states. Although the EU Commission can also try to get direct information from each national regulatory agency, the expert advice given by regulatory networks to the EU Commission might be much more sophisticated and balanced through the internal discussion process within the network. This can increase both the quality of European regulations directly but also lead to better information and awareness about the problems of harmonised regulations due to different problems and conditions in the member states. This can also lead to the recommendation of regulatory solutions that give the national regulatory agencies a larger scope how to apply European rules or even allows for some limited rule-making at the member state level. However, regulatory networks can also help to solve problems of a system, in which rule-making is primarily decentralised. Here a regulatory network can help to give expert knowledge and information to the national regulatory

agencies about the effects of national regulations on other countries, which can influence the national rule-making and solve some of the coordination problems, which usually turn up in the absence of a (strong) centralised rule-making.

Best practices and policy learning: One of the important topics in the economic theory of legal federalism is the potential advantage of regulatory competition in regard to policy innovation, policy learning and diffusion (laboratory federalism). From an evolutionary economics perspective, decentralised regulatory powers allow for parallel experimentation with different regulatory practices, whose positive and negative experiences increase the knowledge about suitable and effective regulatory practices.[6] Even if regulatory competition is only possible as yardstick competition, because a direct choice between different regulations is not allowed, such a parallel experimentation process can lead to a step-by-step improvement of national regulatory practices by mutual learning between the agencies. Networks of regulatory agencies can be very suitable institutions for providing a communication infrastructure and organising a systematic process of the exchange of knowledge and experience, the comparative assessment of regulatory practices, and the spreading of this knowledge for the diffusion of more effective regulatory policies. Therefore the function "best practices and policy learning" is also part of the economic theory of legal federalism and its evolutionary economics perspective on policy innovation and mutual learning (Kerber and Eckardt 2007). Whereas the OMC was organised top-down from the EU Commission without using regulatory networks, benchmarking, the identification of best practices and policy recommendations can also be carried out by the regulatory networks themselves (without the initiative or help of the EU Commission). Therefore regulatory networks can be an instrument of the national regulatory agencies for using yardstick competition in a more effective way in order to further the innovation and diffusion of better regulatory practices. However, this function of regulatory networks can only work permanently, if it is not viewed primarily as a method for achieving more convergence and harmonisation (as this was done in regard to the OMC by the Commission). A permanent process of regulatory innovation, identification of best practices, and diffusion of superior policy is only possible, if also the creation of new variety of regulatory practices is allowed and even encouraged (Kerber and Eckardt 2007, 238–240).

Effective enforcement: From a legal federalism perspective, it need not be optimal that harmonised regulations are also enforced by a European regulatory agency. The advantages of decentralised knowledge (and in the European case also the problem of different languages) will often render a decentralised enforcement of regulations more efficient, even in the case of fully harmonised European regulations. Therefore effective enforcement might need a two-level system of

[6]For laboratory federalism see Oates (1999, 1131–1134); in regard to the interpretation of regulatory competition as an Hayekian evolutionary process of innovation and imitation and linking it to the political science literature on policy innovation and policy learning (e.g., Dolowitz and Marsh 2000), see Kerber and Eckardt (2007, with many references). This evolutionary perspective is close to the small literature in political science about "experimentalist governance" (Sabel and Zeitlin 2008).

enforcement, in which the national regulatory agencies (as well as private enforcement and national courts) might play an important role. However, such a solution might require safeguards for a consistent and equal application of the harmonised rules. Regulatory networks as institutions for exchange of information, communication, and monitoring (both horizontally between the national agencies and vertically in relation to the EU Commission) can facilitate such an effective, equal, and consistent enforcement of regulations, and therefore help to mitigate the problems of decentralised enforcement (see below the example of ECN in Sect. 4). However, this is not limited to the enforcement of harmonised rules. Even in the case of decentralised regulatory powers of the member states, regulatory networks can help to enforce regulations in cases with spillover effects to other member states by facilitating the bi- or multilateral cooperation between national regulatory agencies.

Conflict resolution: In a two-level regulatory regime, in which the regulatory powers in regard to rule-making and/or enforcement are split between a number of different decision-makers and agencies, there might be conflicts between these actors, e.g. in regard to non-clarified delineations of regulatory powers, specific regulatory decisions or the question which regulatory agency should deal with a specific case (case allocation). Regulatory networks can help in different ways. In regard to horizontal conflicts between two regulatory agencies, the discussion of the problems among the experts of the network can facilitate a solution. However, also in regard to the often more difficult vertical conflicts between particular national regulatory agencies and the EU Commission, the regulatory network can try to mediate or even provide arbitration-like functions, either in a purely informal way or in a formalised proceeding (see below the example of BEREC in Sect. 4). Regulatory networks might fulfil an important role in this respect and can therefore help to reduce the costs of conflicts within such two-level regulatory systems.

This discussion has shown that regulatory networks might not only be the result of unsatisfactory political compromises but can also be part of sophisticated optimal solutions for fine-tuning the vertical allocation of regulatory powers in multi-level regulatory systems. This claim requires some qualification but also allows some conclusions: (1) The economic theory of federalism is a normative theory, which analyses what might be optimal. Therefore we do not claim that the existing regulatory networks are already part of an optimal institutional solution. This is a question that has to be analysed for each regulatory network separately. (2) The different trade off problems between centralisation and decentralisation in regard to different regulatory problems imply that regulatory networks (a) might not always be recommendable as part of an optimal solution, and (b) that even if they are, then their optimal institutional design (in regard to memberships, functions, and rights) might be very different. Therefore we cannot expect that a "one-size-fits-all" model for regulatory networks exists. (3) Although the political science literature is right to analyse the evolution of regulatory networks, we claim from a legal federalism perspective that regulatory networks should not primarily be viewed as a transitory phenomenon towards a more centralised and harmonised regulatory system. Rather

regulatory networks should be viewed also as a potentially important part of long-term optimal solutions in multi-level regulatory regimes.

4 Three Case Studies: BEREC, ECN, and ICN

In this section we will take a closer look at three different regulatory networks, the Board of European Regulators for Electronic Communication (BEREC), the European Competition Network (ECN), and the ICN. Since BEREC is a regulatory network for the telecommunication sector and ECN and ICN are regulatory networks of competition authorities, all three regulatory networks have in common that their main objective is the protection of competition. But there are also important differences: Whereas BEREC is active in the field of sector regulation (with natural monopoly problems), ECN and ICN refer to general competition law. A different perspective is offered by the comparison between BEREC and ECN as explicit European regulatory networks with the ICN as a global network of competition authorities.

4.1 BEREC

Since the introduction of full liberalisation in the telecommunication sector in 1998, a comprehensive European regulatory framework was established, leaving a limited scope for own regulatory decisions to the national telecommunication regulators (Haucap and Kuehling 2006). The Framework Directive (2002/21/EC; in short: FD) and in particular the Article 7/7a FD procedure gave the Commission the right to monitor and influence the decisions of the national regulators. Within this regulatory framework and its specific allocation of regulatory powers between the EU Commission and the Member States, BEREC was established as the network of the national regulatory agencies in 2009 (Simpson 2011; Batura 2012). A former plan of the EU Commission for the establishment of a new regulatory agency at the European level failed due to the opposition of the EU Parliament and the national governments (Blauberger and Rittberger 2015, 370–371). BEREC is a fully autonomous Community body with own formal competences and an office in Riga (Latvia). Its decision-making body is the Board of Regulators (composed of representatives of the national regulatory agencies) which decides with a two-thirds majority. Parts of the organisational structure of BEREC are Experts Working Groups, which develop drafts of the network's documents for the Board. The EU Commission is not a member of BEREC but is present as an observer, e.g. in the working groups.

Within the Art. 7/7a FD procedure, which should ensure an effective and equal application of the European rules, BEREC has an own formal role. According to this procedure, the national regulators have to notify the Commission and the other national regulatory agencies of planned decisions in regard to a new market definition, a significant market power of firms or a specific regulatory remedy. If the Commission finds that the intended measure is not compatible with European rules, BEREC is required to analyse the problem and issue an own "opinion" in regard to

Table 1 Analysis of BEREC documents: Initiative, expert advice, soft law, and conflict resolution (May 2011–May 2013; see Table 2 in Appendix)

Who is the initiator?	BEREC	EU Commission	NRA
Number of documents	65	34	1
Role of BEREC?		Yes	No
Expert adviser vis-à-vis the EU level?		51	49
Soft law regulation vis-à-vis the national regulators?		57	43
Dispute resolution?		22	78

this dispute. Whereas in regard to the definition of markets and the assessment of significant market power the EU Commission has the final right to veto a decision of the national regulator (Art. 7 FD), in the case of a remedy it is the national regulator which can make the final decision (Art. 7a FD). In both cases, however, the opinion of BEREC has to be taken into "utmost account" by the Commission or the national regulator.

Before analysing in more depth the functions that BEREC fulfils as regulatory network, we want to present the results of a small empirical study one of us (Julia Wendel) made about the activities of BEREC. Since BEREC does not make decisions, but gives opinions and expert advice, writes reports and issues guidelines, the study focusses on relevant documents, BEREC has published on its website. The time period covered is May 2011 until May 2013. The overall 100 documents include 17 Public Consultations, 39 Reports/Snapshots, 31 Opinions, 4 Guidelines, 6 Common Positions/Approaches, 1 Advice and 2 other documents.[7] The documents were analysed in regard to four questions (for the results see Table 1 and the Appendix):

1. Who initiated the activity? This can be the EU Commission (34 % of the documents, e.g. as part of the Art. 7/7a FD procedure or as queries in regard to specific topics) or a national regulator (1 %), e.g. by asking for technical support. But in 65 % of the cases, BEREC itself took the initiative for making and publishing guidelines, common positions, and reports about certain topics and regulatory questions. This shows a high activity of the network itself (Batura 2012, 6–7).
2. The second question refers to the extent of giving expert advice on rule-making on the European level. This was done in 51 % of the documents.
3. To what extent did BEREC set non-binding rules, standards, and recommendations as part of its soft governance role for the national regulators? In 57 % of the documents BEREC provided guidance to the national regulators.
4. Did BEREC help to solve conflicts? In 22 documents, BEREC was involved in the process of conflict resolution.

[7]Not included are documents, which concern primarily internal organisation issues of the network.

Although BEREC has no formal rule-making power, BEREC contributes a lot to rule-making both at the EU and the member state level. As the empirical results about the published documents show, a very important part of the activities of BEREC is the provision of experts' advice to the rule-making institutions at the EU level. One example is the document BoR (13) 41, which provides a requested opinion by BEREC on a Commission draft on the Recommendation on non-discrimination and costing methodologies. Therefore BEREC could establish itself as a key player for advising the European institutions on telecommunication regulation (Batura 2012, 15). The empirical results also show that BEREC plays an important role in influencing rule-making at the member state level by using its soft governance instruments of developing guidelines and recommendations for the implementation of the European rules by the national regulatory agencies. An example is document BoR (12) 107, which includes legally non-binding Guidelines on the application of Article 3 of the Roaming Regulation. National regulators are expected to consider this document to the utmost account and must state objective reasons for the departure from the Guidelines (BoR (12) 107, 2). This soft governance role of BEREC is directly related to its function of best practices, information distribution and policy learning, because a number of the recommendations and guidelines published by BEREC are based upon the results of working groups for benchmarking and best practices. An example is document BoR (12) 127, which presents a common position on best practice in remedies in a specific market. Therefore Batura (2012, 15) is right to call the use of soft law by BEREC a successful example of "regulation by information" (see also Simpson 2011, 1124).

The objective to establish a functioning internal market in the telecommunication sector is supported by the improvement of effective enforcement of European rules through BEREC by monitoring the regulatory practices of the national regulators and providing channels of information exchange and coordination. The predecessor of BEREC, ERG, has been criticised (and finally replaced) for failing to achieve this goal (Simpson 2011). The monitoring function is well reflected in the network's documents. One example is the report BoR (11) 43 about the implementation of the "Next Generation Access"-Recommendation of the Commission (2010) as key measure of the Digital Agenda.[8] Moreover, with the provision of information channels by BEREC, national decisions might become more sensitive to concerns of other jurisdictions (national and EU ones), and EU decisions might evolve, taking into greater account specific national features (Batura 2012, 15). This can also increase the consistency of European rule application.

The activities of BEREC in regard to conflict resolution did not find much attention in previous research. However, both the legal rules in the Framework Directive and the BEREC documents show that conflict resolution is an important part of the tasks and activities of BEREC. Despite an explicit provision in the Framework Directive for solving horizontal regulatory problems between member

[8]BEREC also provided three opinions on earlier versions of the Recommendation (BoR (11) 43, 6).

states,[9] in the documents only one such case could be found.[10] BEREC is primarily active in regard to vertical conflicts (21 of 22 documents) and this is due to the role of BEREC in the Art. 7/7a FD procedure. If the Commission does not agree with a proposed regulatory measure of the national regulators, then it is a legal requirement that BEREC has to step in and give an own opinion on this dispute. Since the ultimate decision-maker (the Commission in regard to decisions on market definition and significant market power, and the national regulators in regard to remedies) have to take "utmost account" of this opinion, this conflict resolution mechanism falls short of a genuine arbitration solution (with BEREC as arbitrator) but is not far away from it. Thatcher (2011, 803) calls it the "main potential coercive 'power'" of BEREC. An example is document BoR (13) 95, concerning a Spanish case, in which BEREC—after conducting an own separate economic analysis—supports the concerns of the Commission that the Spanish national regulator CMT has not given sufficient evidence for its choice of price market regulation, and therefore recommends that CMT should amend its approach. An analysis of the 21 documents about such vertical conflicts shows that BEREC has agreed in most cases (18 documents) fully or mostly with the concerns of the Commission. However, the approach chosen by the network often differs from the reasoning of the Commission (PWC 2012), which can be interpreted as showing the independence of BEREC from the Commission.

Overall, the analysis of activities and functions of BEREC within the European two-level system of telecommunication regulation supports the claim that this regulatory network helps to optimise the tradeoffs between centralisation and decentralisation. BEREC helps to combine the advantages of decentralised regulatory decision-making due to better knowledge of the specific problems of national markets with the advantages of centralisation in regard to enforcing a consistent application of uniform European rules for achieving a functioning internal market in the telecommunication sector. The role of BEREC as quasi-arbitrator in vertical conflicts is a special characteristic of this regulatory network, which is much less common in other regulatory networks. In this regard BEREC can be seen as helping to balance the advantages and disadvantages of centralised and decentralised decision-making. A recent proposal of the EU Commission, which would include that the Commission also gets a veto right in regard to the remedies of national regulatory agencies, might endanger this balancing role of BEREC, because then the Commission would have in all cases the ultimate decision-making power.[11]

[9]Article 21 FD stipulates that "the competent national regulatory authorities shall coordinate their efforts and shall have the right to consult BEREC in order to bring about a consistent resolution of the dispute".

[10]Document BoR (13) 34 describes a case, where a Belgian company faces a cross-border impediment, which makes a cross-national regulatory action necessary. Ultimately the Dutch regulator (as one of the concerned national regulators) took action and asked BEREC for technical support.

[11]In regard to this proposal and its critique by BEREC, see document BoR (13) 142, 4, and Kerber and Wendel (2014, 190) supporting the rejection of this proposal of the Commission.

This issue is also part of the more general question for the optimal vertical allocation of regulatory powers in the telecommunication sector that cannot be discussed here (see from a legal federalism perspective the thorough analysis of Haucap and Kuehling 2006).

4.2 European Competition Network (ECN)

In contrast to many other European regulatory regimes, there was an early consensus between the EU Commission and the member states that the Single market needs the application of uniform European competition rules, consisting of Art. 101 TFEU (cartel prohibition and exemptions), Art. 102 TFEU (abuse of market dominance), and a common merger policy (Merger Regulation). There was not much resistance against voluntary bottom up-harmonisation of national competition laws with European rules and establishing the principle that the application of national competition laws must not contradict European competition law. Although the European competition law regime still consists of a two-level system of competition laws and competition authorities, it was clear that all relevant regulatory powers are allocated at the EU level.[12] With the ("Modernisation") Regulation 1/2003 the EU Commission started a process of the decentralisation of the application of the European competition rules by allowing both the national competition authorities and the national courts to apply directly Art. 101 and 102 TFEU (Wils 2013). This implied the abolition of the monopoly of the Commission for cartel exemptions according to Art. 101 (3) TFEU. Within this context the European Competition Network was established by the EU Commission as an instrument for ensuring the success of this decentralisation project in regard to the effective and equal application of the European competition rules (Cengiz 2010; Wils 2013).

The European Competition Network consists of the Commission and all national competition authorities in the EU. It is based upon a non-binding "Network Notice" of the Commission, which also has been adopted by the Member States (soft law). It is managed largely by officials of the Commission, and has primarily a hierarchical structure with certain enforcement and monitoring powers of the Commission. The main tasks of the ECN is sharing information, case allocation and ensuring efficient cooperation (Cengiz 2010, 666). Most important is that all competition authorities must inform each other about all cases, in which they apply Art. 101 and 102 TFEU. Between May 2004 and December 2012 the national competition authorities have informed the Commission and other members of the network about 1344 investigations and the intended final decisions in regard to the termination of infringements, imposition of fines, and the acceptance of

[12]The national competition laws as far as they are not fully harmonised can play only a role in small niches of competition law (with the exception of merger policy where the member states still have some scope for smaller mergers which are not subject to EU merger policy).

commitments in 646 cases (Wils 2013, 295). This leads to mutual information between all competition authorities and also allows the Commission to monitor closely the practices of decentralised enforcement. Linked to this top-down monitoring function is the prerogative of the Commission for intervening into the investigations of the national competition authorities, either through soft communication or, in extreme cases, by starting their own investigations. Since the effects of anticompetitive behaviour is often not limited to only one member state, the question which competition authority should deal with a specific case can be crucial for ensuring effective enforcement. Therefore the ECN fulfils an important role in regard to the allocation of cases, both horizontally between the national competition authorities and vertically between the national competition authorities and the Commission.

The literature about the European Competition Network shows clearly that it mainly fulfils the function of supporting effective enforcement (Cengiz 2010; Wils 2013). The mutual sharing of information and monitoring role as well as the allocation of cases are activities of the network that help to ensure an effective, consistent and equal application of European competition rules. In comparison to other networks, the ECN is less active in regard to rule-making both at the EU and member state level, although it also participates in policy discussions, and mutual information and monitoring can lead to a convergence of the practices at the national level. The ECN also has working groups for specific topics, which allow for mutual policy learning. However, benchmarking and best practices do play a smaller role than in other regulatory networks. The ECN also does not provide strong mechanisms for solving conflicts between the competition authorities. The main reason is that the ECN is not needed for conflict resolution, because the Commission has sufficient powers for deciding all conflicts. To what extent can the ECN play an own role in regard to the optimisation of tradeoffs between centralisation and decentralisation in competition policy? Due to the clear decision that the EU Commission as competition authority should have all relevant regulatory powers the ECN cannot play a large independent role and is mostly an instrument of the Commission for ensuring a consistent and effective decentralised enforcement of European competition rules. Therefore Cengiz (2010, 661) is right that the ECN is an atypical example of a European regulatory network. However, it is an interesting and partly surprising result that this hierarchical regulatory network still has been capable of achieving some of the benefits of voluntary, non-hierarchical regulatory networks as, e.g. an extensive communication culture (Blauberger and Rittberger 2015, 372).

From the legal federalism perspective, the ECN can help to reap the advantages of the specific combination of centralized rule-making with decentralised enforcement which characterises the European two-level system of competition laws. Whether this strong harmonisation of competition laws in the EU (and therefore also this hierarchical design of the ECN) is optimal from a legal federalism perspective is, however, an open question. For example, Van den Bergh and Camesasca (2006, 402–446) made a deep and critical analysis of the EU

competition law regime from this legal federalism perspective. Their results show a number of problems of the current system and also convincing arguments against a fully harmonised competition law in the EU. One important line of reasoning emphasises the advantages of decentralised experimentation with diverse competition rules and new regulatory practices for the evolution of an effective competition law. From this perspective, the hierarchical character of the ECN might be seen as a problem. However, it is very interesting that recently competition law scholars have observed that national competition authorities in the EU seem to experiment with new and diverse applications of European competition law, e.g. by developing new case groups or use new enforcement instruments (Monti 2014, 18). Monti raises the question whether the ECN might "evolve into a network that encourages diverse applications of competition law with a view to reflecting on how to best handle certain competition puzzles" (ibid.) but also sees the tension between the hierarchical governance mechanism of the ECN and such an experimentalist approach.

4.3 International Competition Network (ICN)

It is finally interesting to compare these European regulatory networks BEREC and ECN with the ICN, which works as a worldwide network of competition authorities within a very different institutional context (overview: Kovacic and Hollman 2011; Budzinski 2015). In the past all attempts to establish competition law rules at the global level for international markets failed. Therefore competition on international markets can only be protected by national competition law regimes, but this decentralised approach suffers from a number of problems in regard to coordination, conflicts, and particularly effective enforcement. Whether and to what extent the introduction of competition rules and enforcement agencies on the global level can be recommended as part of a multi-level competition law regime, could also be analysed from a legal federalism perspective. Since there are huge obstacles for agreeing on common substantive competition rules on the global level (due to different objectives and conditions in different countries), such analyses suggest that a combination of a more integrated system of procedural rules with minimum standards of substantive competition rules in an otherwise primarily decentralised multi-level competition law regime might be most capable of combining the advantages of centralisation and decentalisation in regard of the protection of competition on international markets (Kerber 2003; Budzinski 2008). However, since it was not possible that the states agree even on basic common rules for competition law, the ICN as an entirely voluntary network of competition authorities was founded in 2001.

In the meantime, the ICN is viewed as a very active and successful regulatory network with 126 members (competition authorities and regulatory agencies) from 111 countries (Sept. 2013) (Kovacic and Hollman 2011). It is a virtual network without an office and a budget, organised by a Steering Group (consisting of

representatives of competition authorities). Its main tasks are convergence, experience-sharing, supporting competition advocacy, and facilitate cooperation (ICN 2011, 4). This has been primarily done by the establishment of working groups, e.g. on cartels, mergers, unilateral conduct, advocacy, and agency effectiveness, who have developed and published best practice recommendations both on substantive as well as procedural rules for competition law and its enforcement. Additionally, the ICN has organised conferences and workshops on specific topics, and is particularly active in the dissemination of the competition experiences and best practices, especially also in regard to emerging and developing countries with new competition laws and often inexperienced competition authorities. Since the best practice recommendations are entirely voluntary, the basic idea of convergence is that states and competition authorities can use them for the enactment of their own competition laws and for competition law enforcement (opt in-solution). Although it is not entirely clear to what extent states and competition authorities have used this possibility, there seems to be a broad consensus that the ICN Recommended Practices and other guidance have influenced the worldwide discussion about competition law and its enforcement.

The ICN differs from the ECN and BEREC in several ways: (1) Since neither competition rules nor a competition authority exist at the global level, the regulatory powers are exclusively allocated at the national levels. Therefore the ICN is a purely voluntary bottom-up project of the national competition authorities. (2) The main function of the ICN is the development of best practice recommendations about the protection of competition and policy learning. (3) Since these best practice recommendations can influence also national policy discussions as well as the practice of national competition authorities, it can also be seen as a soft governance method, which can influence the making of competition rules all over the world. (4) However, the ICN does not monitor the competition law application of the member institutions or help otherwise to increase the effectiveness of competition law enforcement (beyond the provision of best practice recommendations). The ICN, in particular, does not play any role in competition cases, neither through providing mutual information about the cases or supporting directly the cooperation of national competition authorities. (5) Therefore the ICN has also no function in regard to the allocation of cases between national competition authorities (as, e.g. the ECN) nor does it provide any mechanism for solving conflicts between the competition authorities (as, e.g. BEREC).

5 Conclusions

In this article it was shown that networks of regulatory agencies as soft governance instruments can play an important role in multi-level regulatory systems for helping to optimise the tradeoffs between the advantages and problems of centralisation and decentralisation. Therefore regulatory networks can be part of sophisticated solutions for the optimal vertical allocation of regulatory powers in two-level systems of

regulation as in the EU. From the perspective of the economic theory of legal federalism the functions of regulatory networks, which mostly have been discussed already in the political science literature, namely helping rule-making, identifying best practices and promoting policy learning, improving effective enforcement, and supporting conflict resolution can help to combine advantages and avoid problems of centralised and decentralised regulatory powers. Since from a legal federalism perspective, optimal intermediate solutions between centralisation and decentralisation can look very different, it is not surprising that also empirically very different regulatory networks can be observed. This can be seen in the three case studies about BEREC, ECN and ICN. Whereas ECN is a regulatory network in a strongly centralised European regulatory context, ICN operates in an entirely decentralised context. In contrast to both, BEREC works in a regulatory two-level system with still some divided competences. Therefore the different functions of these regulatory networks are not surprising. Important for the further research on regulatory networks is that they should not be viewed primarily as a transitional phenomenon in a final development to centralisation and harmonisation, but should also be seen as potentially important institutions within long-term structures of multi-level regulatory systems.

Appendix

See Table 2.

Table 2 Analysis of the activities of BEREC (published documents. May 2011–May 2013)

Document name	Document type	Who initiates action? 1 = BEREC 2 = COM 3 = NRA	Expert advisor? Yes/No	Regulation by soft law? Yes/No	Dispute resolution? Yes (horizontal/vertical)/ No	Document name	Document type	Who initiates action? 1 = BEREC 2 = COM 3 = NRA	Expert advisor? Yes/No	Regulation by soft law? Yes/No	Dispute resolution? Yes (horizontal/ vertical)/ No
BoR (11)22	2	1	X	–	–	BoR (12)67	1	1	–	X	–
BoR (11)25	2	1	–	X	–	BoR (12)68	1	1	X	–	–
BoR (11)26	2	1	X	X	–	BoR (12)69	3	2	X	–	X(v)
BoR (11)27	2	1	–	x	–	BoR (12)71	3	2	X	–	X(v)
BoR (11)34	2	1	X	X	–	BoR (12)72	3	2	X	–	X(v)
BoR (11)35	2	1	–	X	–	BoR (12)78	2	1	X	–	–
BoR (11)36	2	1	–	X	–	BoR (12)79	2	1	–	X	–
BoR (11)42	2	2	X	–	–	BoR (12)80	2	1	–	X	–
BoR (11)43	2	1	X	–	–	BoR (12)81	2	1	–	X	–
BoR (11)44	1	1	X	X	–	BoR (12)83	5	1	–	X	–
BoR (11)46	2	1	X	–	–	BoR (12)87	1	1	–	X	–
BoR (11)51	2	1	X	X	–	BoR (12)88	5	1	–	X	–
BoR (11)53	2	1	X	X	–	BoR (12)91	3	2	X	–	–
BoR (11)54	1	1	–	X	–	BoR (12)100	1	1	–	X	–
BoR (11)55	2	1	–	X	–	BoR (12)103	5	1	–	X	–
BoR (11)56	2	1	–	X	–	BoR (12)104	1	1	–	X	–
BoR (11)57	2	1	–	X	–	BoR (12)105	1	1	–	X	–
BoR (11)64	3	2	X	–	–	BoR (12)106	2	1	–	X	–
BoR (11)65	3	2	X	–	–	BoR (12)107	4	1	–	X	–
BoR (11)67	4	1	–	X	–	BoR (12)108	2	1	X	–	–
BoR (11)70	1	1	X	X	–	BoR (12)109	3	2	X	–	–

(continued)

Table 2 (continued)

Document name	Document type	Who initiates action? 1 = BEREC 2 = COM 3 = NRA	Expert advisor? Yes/No	Regulation by soft law? Yes/No	Dispute resolution? Yes (horizontal/vertical)/ No	Document name	Document type	Who initiates action? 1 = BEREC 2 = COM 3 = NRA	Expert advisor? Yes/No	Regulation by soft law? Yes/No	Dispute resolution? Yes (horizontal/ vertical)/ No
BoR (11)71	7	2	X	–	–	BoR (12)119	7	1	X	–	–
BoR (11)75	3	2	X	–	X(v)	BoR (12)126	5	1	–	X	–
BoR (11)76	3	2	X	–	X(v)	BoR (12)127	5	1	–	X	–
BoR (12)10	1	1	–	X	–	BoR (12)128	5	1	–	X	–
BoR (12)12	1	1	X	X	–	BoR (12)130	2	1	–	X	–
BoR (12)13	2	1	–	X	–	BoR (12)131	4	1	–	X	–
BoR (12)14	2	1	X	–	–	BoR (12)132	2	1	–	X	–
BoR (12)15	2	1	X	X	–	BoR (12)139	1	1	–	X	–
BoR (12)22	3	2	X	–	X(v)	BoR (12)145	3	2	X	–	–
BoR (12)23	3	2	X	–	X(v)	BoR (12)146	3	1	–	X	–
BoR (12)24	2	1	X	–	–	BoR (13)01	3	2	X	–	X(v)
BoR (12)25	3	2	X	–	–	BoR (13)04	3	4	X	–	X(v)
BoR (12)26	3	2	X	–	X(v)	BoR (13)05	2	1	–	X	–
BoR (12)27	3	2	X	–	X(v)	BoR (13)15	4	1	–	X	–
BoR (12)28	3	2	X	–	X(v)	BoR (13)22	3	2	X	–	–
BoR (12)30	2	2	X	–	–	BoR (13)27	3	2	X	–	–
BoR (12)31	1	2	X	–	–	BoR (13)34	6	3	–	X	X(h)
BoR (12)32	1	1	–	X	–	BoR (13)36	2	1	–	X	–
BoR (12)33	1	1	–	X	–	BoR (13)37	2	1	–	X	–
BoR (12)40	2	1	–	X	–	BoR (13)40	3	2	X	–	X(v)
BoR (12)41	2	1	–	X	–	BoR (13)41	3	1	X	–	–

(continued)

Table 2 (continued)

Document name	Document type	Who initiates action? 1 = BEREC 2 = COM 3 = NRA	Expert advisor? Yes/No	Regulation by soft law? Yes/No	Dispute resolution? Yes (horizontal/vertical)/ No	Document name	Document type	Who initiates action? 1 = BEREC 2 = COM 3 = NRA	Expert advisor? Yes/No	Regulation by soft law? Yes/No	Dispute resolution? Yes (horizontal/ vertical)/ No
BoR (12)51	2	1	–	X	–	BoR (13)47	3	2	X	–	X(v)
BoR (12)52	2	1	–	X	–	BoR (13)53	1	1	–	X	–
BoR (12)53	2	1	–	X	–	BoR (13)54	1	1	–	X	–
BoR (12)54	2	1	–	X	–	BoR (13)55	3	2	X	–	X(v)
BoR (12)55	2	1	–	X	–	BoR (13)73	3	2	X	–	X(v)
BoR (12)56	2	1	–	X	–	BoR (13)93	3	2	X	–	X(v)
BoR (12)61	3	2	X	–	X(v)	BoR (13)94	3	2	X	–	X(v)
BoR (12)66	3	2	X	–	X(v)	BoR (13)95	3	2	X	–	X(v)

In total: 100 documents

Document type 1 = Public Consultation, 2 = Report/Snapshot, 3 = Opinion, 4 = Guideline, 5 = Common Approaches/Positions, 6 = Advice, 7 = Other; BoR = Board of Regulators (Year) document number, X = Yes, – = No, v = vertical; h = horizontal

Source BEREC public document register (http://berec.europa.eu/eng/document_register/welcome/)

References

Arrowsmith, J., K. Sisson, and P. Marginson. 2004. What can "benchmarking" offer the open method of co-ordination? *Journal of European Public Policy* 11: 311–328.

Backhaus, J. 1998. Das Subsidiaritätsprinzip als Baustein einer entstehenden europäischen Verfassung. In *Europäische Integration als ordnungspolitische Gestaltungsaufgabe*, ed. D. Cassel, 79–94. Probleme der Vertiefung und Erweiterung der Europäischen Union, Berlin: Duncker & Humblot.

Batura, O. 2012. Regulatory networks: The role of BEREC in regulation of the common market for electronic communications. 23rd European Regional Conference of the International Telecommunication Society, Vienna, Austria, July 1–4, 2012.

Blauberger, M., and B. Rittberger. 2015. Conceptualizing and theorizing EU regulatory networks. *Regulation and Governance* 9: 367–376.

Borrás, S., and K. Jacobsson. 2004. The Open Method of Co-ordination and new governance Patterns in the EU. *Journal of European Public Policy* 11(2): 185–208.

Börzel, T. 2010. European Governance: Negotiation and Competition in the Shadow of Hierarchy. *Journal of Common Market Studies* 48(2): 191–219.

Budzinski, O. 2008. *The Governance of Global Competition: Competence Allocation in an International Multilevel Competition Policy System*. Cheltenham: Edward Elgar.

Budzinski, O. 2015. International Antitrust Institutions. In *The Oxford Handbook of International Antitrust Economics*, edited by R.D. Blair and D.D. Sokol, vol. 1, 119–146. Oxford: Oxford University Press.

Cengiz, F. 2010. Multi-level governance in competition policy: The European Competition Network. *European Law Review* 35(5): 660–677.

Coen, D., and M. Thatcher. 2008. Network Governance and Multi-level Delegation: European Networks of Regulatory Agencies. *International Public Policy* 28(1): 49–71.

Danielsen, O., and K. Yesilkagut. 2014. The Effects of European Regulatory Networks on the Bureaucratic Autonomy of National Regulatory Authorities. *Public Organizational Review* 14: 353–371.

Dehousse, R. 1997. Regulation by Networks in the European Community: The Role of European Agencies. *Journal of European Public Policy* 4(2): 246–261.

Dolowitz, D., and D. Marsh. 2000. Learning from abroad: The role of policy transfer in contemporary policy making. *Governance* 13: 1–24.

Eberlein, B., and E. Grande. 2005. Beyond delegation: Transnational regulatory regimes and the EU regulatory state. *Journal of European Public Policy* 12(1): 89–112.

Eberlein, B., and A. Newman. 2008. Escaping the International Governance Dilemma? *Incorporated Transgovernmental Networks in the European Union, Governance: An International Journal of Policy, Administration, and Institutions* 21(1): 25–52.

Esty, D.C., and D. Geradin eds. 2001. *Regulatory Competition and Economic Integration: Comparative Perspectives*. Oxford: Oxford University Press.

Garcimartín, F.J. 1999. Regulatory competition: A private international law approach'. *European Journal of Law and Economics* 8: 251–270.

Gilardi, F. 2012. Transnational diffusion: Norms, ideas, and policies. In *Handbook of International Relations.* Edited by W. Carlsnaes, T.T. Risse and B. Simmons, 453–477. Thousand Oaks: Sage Publications.

Haucap, J., and J. Kuehling. 2006. Eine effiziente vertikale Verteilung der Exekutivkompetenzen bei der Regulierung von Telekommunikationsmärkten. *Zeitschrift für Wirtschaftspolitik* 55: 324–356.

Heine, K., and W. Kerber. 2002. European corporate laws, regulatory competition and path dependence. *European Journal of Law and Economics* 13: 43–71.

Héritier, A. 2003. New Modes of Governance in Europe: Increasing Political Capacity and Policy Effectiveness? In *The State of the European Union, 6—Law, Politics and Society*. Edited by T. A. Börzel and R. Cichowski. Oxford: Oxford University Press.

ICN. 2011. The ICN's vision for its second decade. Presentation at the 10th annual conference of the ICN. Available at http://www.internationalcompetitionnetwork.org/uploads/library/doc755.pdf.

Kerber, W. 2003. An International Multi-Level System of Competition Laws: Federalism in Antitrust. In *The Future of Transnational Antitrust—From Comparative to Common Competition Law*, ed. J. Drexl, 269–300. Staempfli/Kluwer: Berne/The Hague/London/Boston.

Kerber, W. 2008. European system of private laws: An economic perspective. In The Making of European Private Law. Edited by F. Cafaggi and H. Muir Watt, 64–97. Cheltenham: Edward Elgar.

Kerber, W., and M. Eckardt. 2007. Policy learning in Europe: The open method of co-ordination and laboratory federalism. *Journal of European Public Policy* 14(2): 227–247.

Kerber, W., and S. Grundmann. 2006. An optional European contract law code: Advantages and disadvantages. *European Journal of Law and Economics* 21(3): 215–236.

Kerber, W., and Wendel, J. 2014. Regulation of network sectors in the EU: A federalist perspective, Wirtschaftspolitische Blätter, 181–193.

Kirchner, C. 1997. Competence Catalogues and the Principle of Subsidiarity in a European Constitution. *Constitutional Political Economy* 8: 71–87.

Kovacic, W.E., and H. Hollman. 2011. The International Competition Network: Its Past, Current, and Future Role. *Minnesota Journal of International Law* 20: 274–325.

Lavrijssen, S., and L. Hancher. 2008. Networks on track: From European regulatory networks to European regulatory network agencies. *Legal Issues of Economic Integration* 34(1): 23–55.

Levi-Faur, D. 2011. Regulatory networks and regulatory agencification: Towards a Single European Regulatory Space. *Journal of European Public Policy* 18(6): 810–829.

Maggetti, M. 2014. The rewards of cooperation: The effects of membership in European regulatory networks. *European Journal of Political Research* 53: 480–499.

Maggetti, M., and F. Gilardi. 2011. The policy-making structure of European regulatory networks and the domestic adoption of standards. *Journal of European Public Policy* 18(6): 830–847.

Marciano, A., and J.-M. Josselin eds. 2002. *The Economics of Harmonizing European Law*. Cheltenham: Edward Elgar.

Marciano, A., and J.-M. Josselin eds. 2003. *From Economic to Legal Competition*. New Perspectives on Law and Institutions in Europe, Cheltenham: Edward Elgar.

Marks, G., L. Hooghe, and K. Blank. 1996. European integration from the 1980s: State-centric v. multi-level governance. *Journal of Common Market Studies* 34: 341–378.

Monti, G. 2014. Independence, Interdependence and Legitimacy: The EU Commission, National Competition Authorities, and the European Competition Network, EUI Working Papers LAW 2014/01.

Oates, W.E. 1999. An Essay on Fiscal Federalism. *Journal of Economic Literature* 37: 1120–1149.

Pelkmans, J. 2006. *European Integration. Methods and Economic Analysis*, 3rd ed. Harlow et al.: Prentice Hall.

Powell, W. 1990. Neither Market nor Hierarchy: Network Forms of Organization. *Research in Organizational Behavior* 12: 295–336.

PriceWaterhouseCooper. 2012. Study on the Evaluation of BEREC and the BEREC Office Workshop Presentation, Brussels, October 8, 2012. Available at http://ec.europa.eu/information_society/newsroom/cf/dae/document.cfm?doc_id=1010.

Radaelli, C.M. 2008. Europeanization, Policy Learning, and New Modes of Governance. *Journal of Comparative Policy Analysis: Research and Practice* 10(3): 239–254.

Sabel, C.F., and J. Zeitlin. 2008. Learning from Difference: The New Architecture of Experimentalist Governance in the EU. *European Law Journal* 14(3): 271–327.

Sandström, A., and L. Carlsen. 2008. The Performance of Policy Networks: The Relation between Network Structure and Network Performance. *The Policy Studies Journal* 36(4): 497–524.

Saz-Carranza, A., and F. Longo. 2012. The Evolution of the governance of regulatory networks: The case of the European Telecommunications regulatory network, Jerusalem Papers in Regulation & Governance, Working Paper No. 41.

Simpson, S. 2011. 'New' Governance in European Union Policy Making: Policy Innovation or Political Compromise in European Telecommunications? *West European Politics* 34(5): 1114–1133.

Slaughter, A. 2005. *A new world order*. Princeton: Princeton University Press Group.

Sun, J.-M., and J. Pelkmans. 1995. Regulatory Competition in the Single Market. *Journal of Common Market Studies* 33: 67–89.

Thatcher, M. 2011. The creation of European regulatory agencies and its limits: A comparative analysis of European delegation. *Journal of European Public Policy* 18(6): 790–809.

Van den Bergh, R. 2000. Towards an Institutional Legal Framework for Regulatory Competition in Europe. *Kyklos* 53: 435–466.

Van den Bergh, R., and P.D. Camesasca. 2006. *European Competition Law and Economics. A Comparative Perspective*, 2nd ed. London: Sweet & Maxwell.

Wils, W. 2013. Ten Years of Regulation 1/2003—A Retrospective. *Journal of European Competition Law and Practice* 4(4): 293–301.

Zeitlin, J. 2005. The open method of coordination in action: Theoretical promise, empirical realities, reform strategy. In *The Open Method of Coordination in Action: The European Employment and Social Inclusion Strategies,* Edited by J. Zeitlin, P. Pochet and M. Magnusson, 441–498. Brussels: P.I.E.-Peter Lang.

Two Treatments of Pluralism: Canada and the United States

Margaret F. Brinig

Most of my interactions with Jüergen Backhaus over the past 20 years stemmed from his editorship of the European Journal of Law and Economics. Some of these involved articles that were for some reason too controversial to find ready acceptance elsewhere. For example, one, coauthored with economist Michael Alexeev was entitled Fraud in Courtship: Annulment and Divorce,[1] and maintained that there is an optimal amount of premarital fraud. Even more provocative, perhaps, was a sociobiology piece with Douglas Allen, Sex, Property Rights and Divorce,[2] providing a theoretical and empirical explanation for the "7-year itch." Perhaps after taking these chances on me he felt he could ask me to write for him as well as to referee for the journal, as I do quite regularly. At Jüergen's invitation, I therefore responded to a piece bemoaning elimination of a law requiring equality of housework caused by the reunification of Germany.[3] Finally, I had the pleasure of his soliciting a chapter from me on family law for the Elgar Companion.[4] So it is somewhat in the spirit of this rather inventive exchange that I submit this contribution to the collection celebrating him.

Fritz Duda Family Professor Of Law, University of Notre Dame. Thanks are due to Hazel Thompson-Ahye and the Caribbean Regional Meeting of the International Society for Family Law as well as attendees of my keynote address at the Canadian Law and Economics Association meeting in 2010.

[1]Margaret F. Brinig and Michael V. Alexeev, Fraud in Courtship: Annulment and Divorce, 2 European Journal of Law and Economics 45–63 (1995).
[2]Douglas W. Allen and Margaret F. Brinig, Sex, Property Rights and Divorce, 5 European Journal of Law and Economics 211 (1998).
[3]Margaret F. Brinig, Equality and Sharing: Views of Households Across the Iron Curtain, 7 European Journal of Law and Economics 55 (1998).
[4]Margaret F. Brinig, Family Law, Chap. 4 of the Elgar Companion to Law and Economics (Jüergen Backhaus, ed., 2003).

M.F. Brinig (✉)
School of Law, University of Notre Dame, Notre Dame, IN, USA
e-mail: Margaret.brinig.1@nd.edu

A. Marciano and G.B. Ramello (eds.), *Law and Economics in Europe and the U.S.*,
The European Heritage in Economics and the Social Sciences,
DOI 10.1007/978-3-319-47471-7_8

Canada and the United States, while similar in many ways, diverge substantially when it comes to family law. Canada's marriage and divorce law is national, while the U.S. family law is largely governed by state law. This makes rules in the United States heterogeneous compared to those in Canada, and thus easier to tailor to the preferences of people living in the various states.

More important for this paper, Canada's approach to pluralism, dealing with nontraditional family forms, differs as well. In C-23, the Modernization of Benefits Act, Canada gave unmarried couples (and their children) the same federal benefits and obligations as to married couples.[5] Canada now recognizes same-sex marriages as well as granting many benefits to heterosexual couples who do not marry.[6] These legal changes were also reflected in Canadian writing on the family. In 2001, the Canadian Law Commission after much study released a report called *Beyond Conjugality*, which included the words, "The state cannot create healthy relationships; it can only seek to foster the conditions in which close personal relationships that are reasonably equal, mutually committed, respectful and safe can flourish."[7]

In contrast, states in the United States consistently maintain differences between married and unmarried couples,[8] and the federal government has enacted legislation favoring marriage and confining it to a man and a woman (while allowing states to do so).[9] The legislation, as in Canada, is reflected in both academic studies and political documents. For example, in the same year as *Beyond Conjugality,* University of Chicago demographer Linda Waite published her much discussed *The*

[5]Modernization of Benefits and Obligations Act (S.C. 2000, c. 12). However, the 2013 case of Quebec (Attorney General) v. A., 2013 SCC 5 (2013), in a 5-4 decision, allowed Quebec to maintain its own separate status for what it calls "de facto" couples. While they are allowed some relief under restitutionary principles in egregious situations, and while child custody and support are treated the same way as for dissolving marriages, most property and support regimes are not. They do have access to the substantial federal benefits provided all couples by C-23, and of course may contract between themselves. The majority grounded its reasoning on the decision to recognize the partners' autonomy and, in part, upon Quebec's formal legal equality between the sexes.

[6]C-38, An Act respecting certain aspects of legal capacity for marriage for civil purposes, 38th Parliament - 1st Session (2005). The United States has been recognizing such marriages piecemeal, especially following the Supreme Court case of United States v. Windsor, 670 US—, 133 S.Ct. 2675 (2013).

[7]Id. at xxiii.

[8]The closest exception that takes into account heterosexual relationships is Washington's "meretricious relationships" law. A summary of the law treating unmarried couples in the United States can be found in the American Law Institute, Principles of Family Dissolution, Reporter's Notes to Chap. 6: Analysis and Recommendations (2002), at 914–16, and Comments to § 6.03, id. at 918–19; as well as in Ira Ellman, Paul Kurtz & Elizabeth Scott, Family Law: Cases, Text, Problems 919–82 (5th ed. 2009).

[9]The Defense of Marriage Act, Pub.L. 104–199, 110 Stat. 2419, enacted September 21, 1996, 1 U. S.C. § 7 and 28 U.S.C. § 1738C. The Act was overruled by Obergefell v. Hodges, 576 U.S. – (2015). At least in the United States, recognizing same-sex marriages does not necessarily reflect an adoption of the policy that marriage does not matter. See, e.g., Baskin v. Bogan, 766 F.3d 648, 658–59, 663 (7th Cir. 2014)(stressing advantages of marriage recognition for adopted children).

Case for Marriage: Why Married People are Happier, Healthier, and Better Off Financially.[10] President Obama, while taking a progressive stance on national health care, has also touted marriage in his *Audacity of Hope*[11] (2006).

> Finally, preliminary research shows that marriage education workshops can make a real difference in helping married couples stay together and in encouraging unmarried couples who are living together to form a more lasting bond. Expanding access to such services to low-income couples, perhaps in concert with job training and placement, medical coverage, and other services already available, should be something everybody can agree on ….[12]

In summary, the difference between the two treatments is that Canada supports a diversity of relationships positively (through providing for financial assistance and legal recognition) and through its public policy. The United States, while tolerating most family forms,[13] formally recognizes only marriage and adoption, leaving adults in heterodox relationships to private support or contract.

In both these North American jurisdictions, people live in a variety of family forms. While most heterosexual couples marry, some do not. This paper considers the effects of the differing policies on young people in two minority groups, the Québécois in Canada and African-Americans in the United States, both of which groups de facto eschew formal marriage.[14] Both are relatively impoverished groups, and both historically have suffered discrimination and been underrepresented among the power elites. Yet despite these surface differences, the two groups diverge in terms of the mental health of their youth, and quite notably in terms of the rate at which they commit suicide. This paper will attempt to portray these similarities and differences as well as propose several reasons for the differing results.

---

[10](New York: Broadway, 2001). As the title implies, Waite presented studies showing that married couples, holding other sociodemographic factors constant, perform better than do their single counterparts.

[11]The Audacity of Hope: Thoughts on Reclaiming the American Dream (Crown Pub. New York, 2006).

[12]Id. at 334.

[13]Lawrence v. Texas, 539 U.S. 558 (2003), disallowed criminalization of sexual activities between consenting adults that do not harm others. The Court wrote this would not include bigamous, polygamous, or incestuous relationships, nor would its analysis require same-sex marriage. Obergefell v. Hodges, 576 U.S. – (2015) held that same-sex couples had a constitutional right to marry.

[14]As with all such statistical and demographic studies, there are of course exceptions. Some Québécois and African-Americans do marry; some are wealthy; some have reached the apex of power in their countries (such as Pierre Trudeau and Jean Chrétien in Canada and of course Barack Obama in the United States), and the vast majority of their youth do not commit suicide. On the other side of the coin, some wealthy, married, and privileged Americans and Canadians produce youth with problems, including suicide.

1 A Portrait of the Family in Contemporary Quebec and for African-Americans

While a minority of Canadians[15] are identified as Québécois,[16] Quebec contains a disproportionate percentage of the cohabiting couples and single-parent families in Canada.[17] In 2009, nearly half the couples in the province were unmarried,[18] and more than half the births in Quebec were to unmarried mothers.[19] About 31.5 % of households live in "common law" unions compared to 15.7 % for all of Canada counting Quebec, or 9.2 % without it.[20] In Quebec, as in other provinces and in the United States, the United States, cohabiting relationships are only half as stable as are marital ones (Table 1).[21]

Nor has this difference in instability changed over the years, even though the percentage of cohabiting couples in the province is allegedly the highest in the world.[22] Quebec also boasts a higher divorce rate than the other provinces in Canada, 49.9 % by age 50 (Fig. 1 and Table 2).[23]

---

[15]That is, about 23 % according to the 2014 Canadian Census. See Statistics Canada Summary Tables, Population by marital status and sex, by province and territory (Quebec, Ontario, Manitoba, Saskatchewan), available at http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/famil01b-eng.htm (March 24, 2015). This shows the population of Canada as 35,540,419, with Quebec's as 8,214,672.

[16]This group is defined for purposes of this paper as those who live in Quebec, speak French, and are white. It therefore does not include Caribbean immigrants to Montreal, who may also speak French, nor French Canadians living in Newfoundland or British Columbia (Arcadians), nor the small Anglophone population (about 10 %) who live in the province. The linguistic breakdown is available at http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/demo11b-eng.htm (March 24, 2015).

[17]See generally Zheng Wu, Economic Circumstances and the Stability of Nonmarital Cohabitation, Cat. 9870, available at http://www.statcan.gc.ca/pub/75f0002m/75f0002m1998010-eng.pdf., at 20. For the relative numbers, see http://www.imfcanada.org/issues/canadian-families-global-context (last visited March 24, 2015). The percentage living in common law couples in Quebec is 31.5 %, and unmarried 16.6 %. For Canada not including Quebec, the percentage living in common law couples is 9.2 %, and unmarried 3.3 %.

[18]47.3 %, adding together the two categories of unmarried in footnote 18 found in ZHeng Wu.

[19]59.3 %, according to Statistics Canada, Table 102-4506, 2009.

[20]http://www12.statcan.gc.ca/census-recensement/2011/as-sa/98-312-x/2011001/tbl/tbl2-eng.cfm (computations are mine, March 24, 2015).

[21]France-Pascale Menard, What Makes It Fall Apart? The Determinants of the Dissolution of Marriages and Common Law Unions in Canada, 2 McGill Soc. Rev. 59, 68 & Fig. 1 (2011).

[22]Dana Hamplová, Céline Le Bourdais and Évelyne Lapierre-Adamcyk, Is the Cohabitation-Marriage Gap in Mondey Pooling Universal? At 27 and Table 2, available at http://iussp.org/sites/default/files/event_call_for_papers/Money%20management%20(Hamplova%20et%20al).pdf. The assertion about the world is made id at page 3, attributed to Statistics Canada, 2012.

[23]Vanier Family Institute, Oct. 26, 2011, Fascinating Families, Four in Ten Marriages End in Divorce.

Table 1 Probability for Women to Separate, by Type of First Union, Quebec 2006

	Quebec		Other provinces	
	50–59 years	30–39 years	50–59 years	30–39 years
Probability for women to go through at least one separation According to whether the first union was …	33.8	45.8	30.5	40.6
Marriage	30.6	26.8	30.2	30.7
Common law	64.8	55.3	60.4	66.3

General Social Survey—Cycle 15—Changing Conjugal Life in Canada, at 9 and Table 1 (2002)
Source General Social Survey, Statistics Canada, 2006

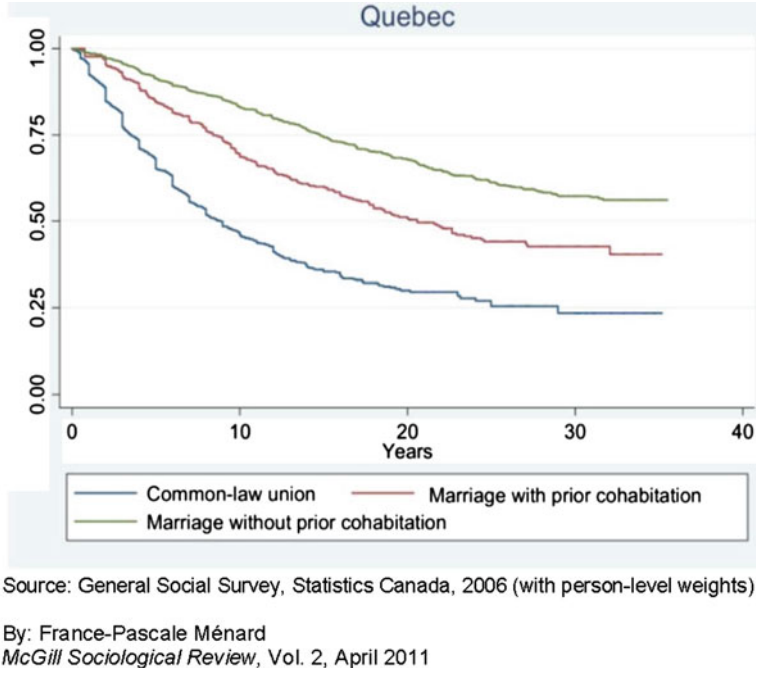


Fig. 1 Union dissolution in Quebec, by union type (time-varying)

As previously indicated, the story for African-Americans looks much the same. In 2010, 25 % of African-American, or Black, women over 35 had never married (compared to 7 % for white women),[24] and in 2008, 71.8 % of all births in this

[24]Diana B. Elliott, Kristy Krivakas, Matthew W. Brault and Rose M. Kreider, Historical Marriage Trends from 1890–2010: A Focus on Race Differences, SEHSD Working Paper Number 2012-12 and Fig. 5, available at http://www.census.gov/hhes/socdemo/marriage/data/acs/ElliottetalPAA2012paper.pdf (March 23, 2015)

Table 2 Cumulative percentages of separation 12 years after the beginning of the union, according to union type and cohort, Quebec and other Canadian provinces

Union cohort	Cohabiting union[a]		Direct marriage	
	Quebec	Other Canadian provinces	Quebec	Other Canadian provinces
1970–1979	41.6	47.8	13.8	16.4
1980–1989	47.0	44.6	18.8	17.5
1990–1999	49.2	46.1	24.7	16.2

Source Life tables derived from statistics Canada, 2006 General Social Survey on Family Transactions, cycle 20, Public Use Microdata files

[a]Union started as a cohabitation, transformed or not into a marriage

group occurred outside marriage.[25] While the single-mother-headed family has occurred for some time among the African-American population, as in Quebec, it is far from stable. Even at the end of 3 years, couples remain together only slightly more than half the time, while at the end of 5 years, the number of intact relationships has declined to only 26 % (Figs. 2 and 3).[26]

While short-lived relationships are painful for those involved when they end, what is more important is that children in them will experience disruption in their living patterns. This holds true both in "common law" families in Quebec and among cohabiting African-Americans in the United States (Figs. 4 and 5).[27]

How old is the child likely to be when the parents separate?

Again, among African-Americans, the result is similar. According to a study based on the National Survey of Family Growth, three-fifths of Black children will no longer be living with both parents at age 5.[28] This is more than twice as high a

[25]6.5 million couples of all races cohabited in the United States in 2007. America's Families and Living Arrangements: 2007. The last figures for births appeared in the 2012 Statistical Abstract, Tables 80 and 85. http://www.census.gov/compendia/statab/cats/births_deaths_marriages_divorces.html (March 23, 2015).

[26]Marriage and Cohabitation in the United States, Fig. 15, based on the National Survey of Family Growth, Cycle 6 (2002)(including childless couples), Table 18. See comparable numbers from the relatively poor, urban families in the Fragile Families Study, at Robert A. Hummer and Erin R. Hamilton Race and Ethnicity in Fragile Families, 20 Future of Children 113, 119 and Fig. 4 (2010) (showing higher rates of breaking up and lower rates of staying married for African-American couples at 3 and 5 years after the birth of a child, when parents cohabited at the child's birth). The Fragile Family disruptions are less likely than for childless couples, for "Cohabiting unions in which children are born tend to last longer than those that are childless, but they still remain significantly more unstable than marriages." Valerie Martin, Céline Le Bourdais & Evelyne Lapierre-Adamcyk, Stepfamily instability in Canada—The impact of family composition and union type, 23 Journal of Family Research 196, 197 (2011).

[27]See, e.g., Sara McLanahan and Audrey Berk, Parental Relationships in Fragile Families, 20 The Future of Children 17 (2010).

[28]Wendy D. Manning, Pamela Smock and Deborun Majumdar, The Relative Stability of Cohabiting and Marital Unions for Children, 23 Population Research and Policy Review 135, 146 (2004). Note that the terms Black and African-American are used interchangeably in this piece, as they are in the literature and on government websites.

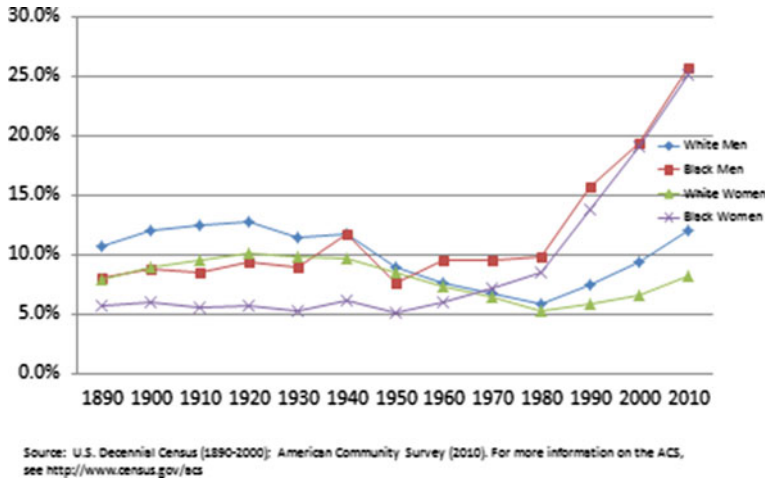


Fig. 2 Never-married among Blacks, United States

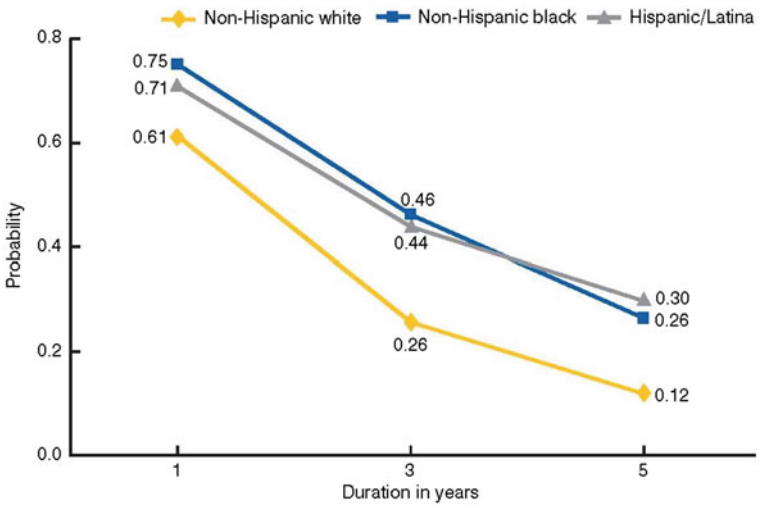


Fig. 3 Probability of separation: US by race [Vital and Health Statistics (U.S.), Marriage and Divorce in the United States: A Statistical Portrait Based on Cycle 6 (2002) of the National Survey of Family Growth, Series 23, No. 28, February, 2010, Center for Disease Control, Atlanta, 2010, page 9 & Fig. 15]

probability of disruption in their living situation as for those born to married parents, even controlling for other factors.[29]

[29]Id. at 148 and Table 2.

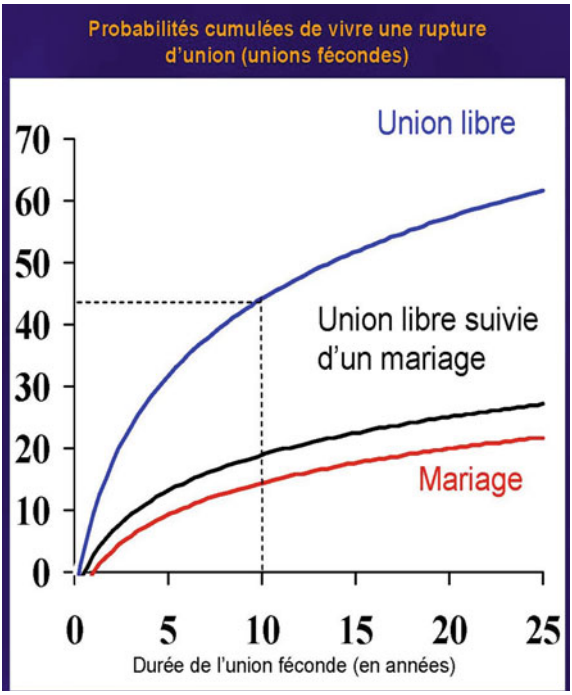


Fig. 4 Cumulative probability of a child's living through a disrupted union

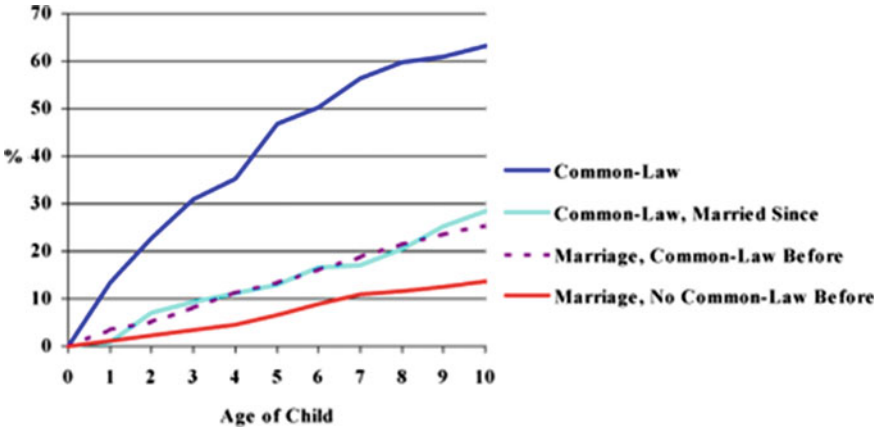


Fig. 5 Likelihood of disruption of various forms of union by child's age

1.1 *Adolescent Outcomes in African-American and Québécois Families*

In my prior work, I have noted that despite material disadvantages and lower educational attainment, Black adolescents do remarkably well from a psychological standpoint.[30] They display no more depression or anxiety, less substance abuse, and no more delinquency than other Americans once income is taken into account.[31] Furthermore, as we will see shortly, they remain optimistic about the future compared to their peers.

On the other hand, French-speaking young people in Quebec have the highest provincial suicide rate in Canada,[32] and one of the highest in the Western world. They are more depressed,[33] and less optimistic than other adolescents, as I will discuss below. They abuse alcohol at a higher rate than do most Canadian adolescents.[34] Figures 6 and 7 show the suicide rate, one in terms of its change over time, one in comparison to the rest of Canada. Figure 6 shows that the rate has

[30]Margaret F. Brinig and Steven L. Nock, The One Size Fits All Family, 49 Santa Clara Law Review 137 (2009).

[31]These results are based on regressions from the Panel Survey of Income Dynamics (PSID), Child Development Supplement CDS), 2002–03. Some are reported Id. at 146–47 (text) and 163, Figs. 2 and 3. Others appear in Margaret F. Brinig and Steven L. Nock, Legal Status and Effects on Children, 5 University of St. Thomas Law Journal 548, 579 and Table 10 (2007); or Margaret F. Brinig and Steven L. Nock, How Much Does Legal Status Matter? Adoptions by Kin Caregivers, 36 Family Law Quarterly 449, 473–74 and Tables 2 and 3 (2002–03).

[32]Andrea Shaver, Teen Suicide, Statistics Canada BP-236E, available at http://dsp-psd.pwgsc.gc.ca/Collection-R/LoPBdP/BP/bp236-e.htm (last visited April 21, 2011). An update does not include a graph but shows that the comparable figures for Canada as a whole in 2011 were, for males 15–24, 8.42 per 100,000 and 42.63 per 100,000 for males 15–24 in Quebec. The population figures for the province were calculated from Population of Quebec, 1971–2014, http://www.stat.gouv.qc.ca/statistiques/population-demographie/structure/index_an.html (March 24, 2015). The suicide rates come from **suicide**prevention.ca/…/**2014/**…/**Suicide**-**Rate**-Across-**Canada**-and Provinces 10-14-14.xlsx.

[33]Amy H. Cheung and Carolyn Dewa, Canadian Community Health Survey: Major Depressive Disorder and Suicidality in Adolescents, 2 Healthy Policy 76, 82 and Table 1 (2006).

[34]See, e.g., Mark Zoccolillo, Frank Vitaro and Richard E. Tremblay, Problem Drug and Alcohol Use in a Community Sample of Adolescents, 38 Journal of the American Academy of Child and Adolescent Psychiatry 900 (1999), showing in a survey of adolescents in Quebec, 62.2 % had drunk alcohol more than given times in their lifetime, and 50 % of both boys and girls drank at least once a week. Id. at 902 and Table 2, while one-third of the boys and a quarter of the girls drank alcohol in the morning.

In 2008, in a survey of 7–12 graders, the proportion of those reported drinking at least once a month in the previous year was highest in Quebec at 28.5 %, followed by BC at 26.0 %, Atlantic at 18.5 %, Ontario at 16.6 % and lowest in the Prairies at 15.7 %. Quebec also had the highest reported use of tobacco at 42.0 % while Ontario had the lowest at 16.4 %. David Hammond et al., Illicit Drug Use Among Canadian Youth Revue Canadienne de Santé Publique Vol. 102, No. 1, at 10 and Table 4, available at http://davidhammond.ca/wp-content/uploads/2014/12/2011-CJPH-Youth-Substance-Use-Hammond1.pdf.

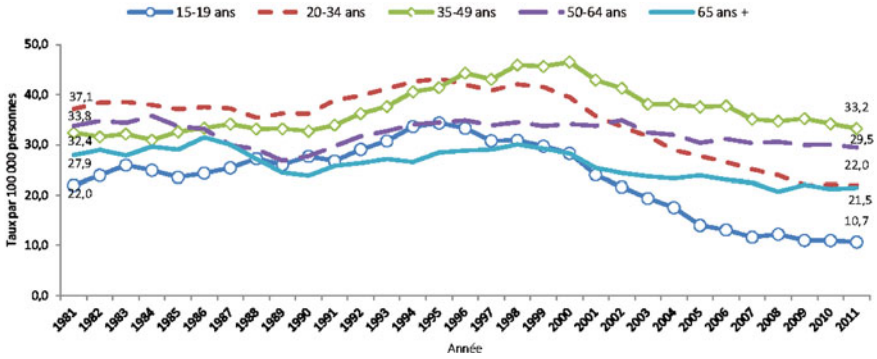


Fig. 6 Taux de mortalité par suicide selon les groupes d'âge, hommes, ensemble du Québec, 1981 à 2007

decreased from a peak in about 1995 of 34.4/100,000 for males 15–24 (though it is higher than that today), but remains comparatively higher than other provinces.[35] The second shows a comparison for all ages over years 2005–2011.[36]

1.2 Reasons for the Difference

A Montreal psychiatrist,[37] citing the Quiet Revolution, when the Church became uninvolved with governmental functions, as well as family breakdown (increased failure to marry), argues that the social upheaval in Québec since the 1960s has affected troubled teenagers by giving them nothing stable to fall back on. Further, he stated, "We are a society that values the quality of life rather than its quantity… Life is [seen by some teenagers as] not worth living if you cannot guarantee its quality."

[35]Id. at 83 and Table 2 (2006).

[36]Although some statistics data is available for each province, the age breakdown varies and some do not collect it each year. The statistics in Excel form compiled by the Centre for Suicide Prevention can be downloaded from https://suicideinfo.ca/Library/AboutSuicide/Statistics.aspx. Population data for the relevant ages was downloaded from http://www5.statcan.gc.ca/cansim/a26 to calculate the rates.

[37]Mounir Samy, Montreal General Hospital, in Teen Suicide (BP-236E) http://dsp-psd.pwgsc.gc.ca/Collection-R/LoPBdP/BP/bp236-e.htm 4 of 9.

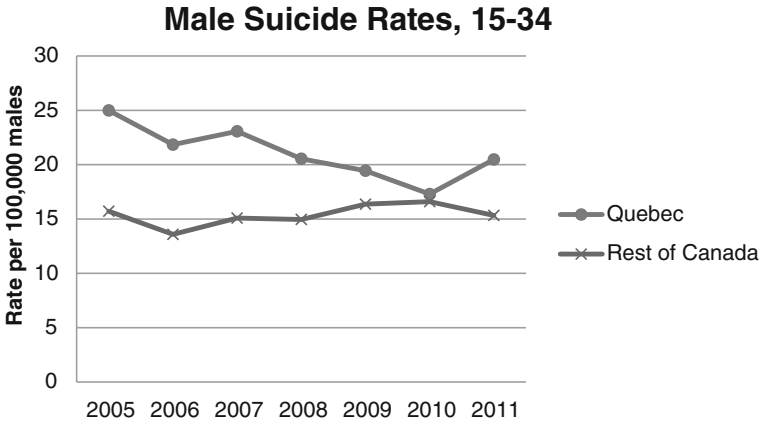


Fig. 7 Suicide Races for Youth, Canada and Quebec

This hypothesis is consistent with my own. One of the central features of my recent book, *Family, Law, and Community*,[38] is that typically families need community support to flourish. This support may come from formal legal status, such as marriage or adoption. It may also stem from mediating institutions, such as religious organizations, parochial schools,[39] or perhaps military service.

What I have reported above for suicide among youths also holds true for their depression[40] and alcohol use.

The tables that follow consider smoking, comparing Quebecois, Canadian, and other provincial populations,[41] followed by similar data involving the United States as a whole compared to its African-Americans (Fig. 8 and Table 3).

There are similar differences in alcohol use and in binge drinking (Figs. 9 and 10).

1.3 Optimism Among African-Americans and the Québécois, a Provisional Study

Another way of looking at the difference is to consider what might be opposite (positive) outcomes related to optimism. For this paper, I have compared data from

[38]Margaret F. Brinig: Family, Law and Community: Supporting the Covenant (University of Chicago Press, 2010).

[39]See Margaret F. Brinig and Nicole Garnett, Catholic Schools and Broken Windows, 9 Journal of Empirical Legal Studies 347 (2012).

[40]Amy Cheung, footnote 33 and Table 1 (2002 data).

[41]Canadian Tobacco Use Monitoring Survey, http://www.hc-sc.gc.ca/hc-ps/tobac-tabac/research-recherche/stat/_ctums-esutc_prevalence/prevalence-eng.php (1999).

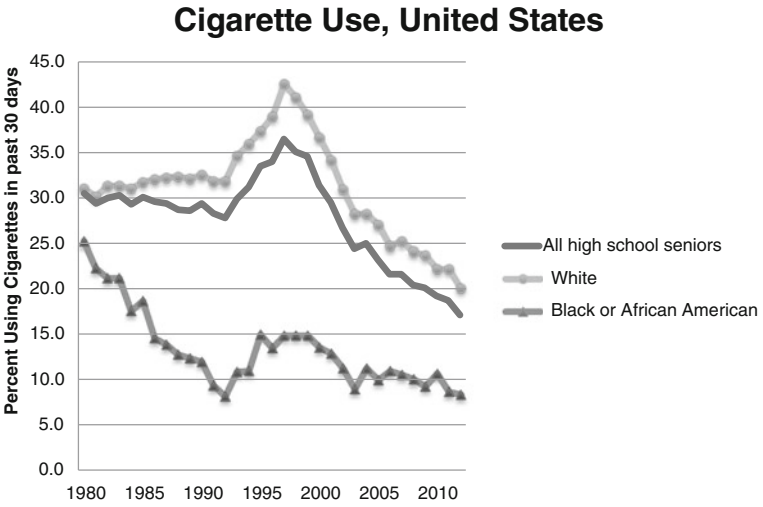


Fig. 8 U.S. Smoking Among High School Students [Chart produced from Excel version of Table 61, Use of selected substances in the past 30 days among high school seniors, 10th graders, and 8th graders, by sex and race: United States, selected years 1980–2012. National Institutes of Health, National Institute on Drug Abuse, Monitoring the Future Study, annual surveys. http://www.cdc.gov/nchs/hus/contents2013.htm#061]

Table 3 Smoking in Canada and by Province, 1999

Province	Age group				
	15+ (%)	15–19 (%)	20–24 (%)	15–24 (%)	25+ (%)
Canada	25	28	34	31	24
NFLD	28	30	36	33	26
PEI	27	27	40	33	26
NS	29*	30	36	33	28
NB	28	26	39	33	27
Que	28	36	39	38	26
Ont	24	24	30	27	23
Man	23	30	35	32	21
Sask	25	34	36	35	23
Alb	27	25	42	33	26
BC	20	23	26	25	19

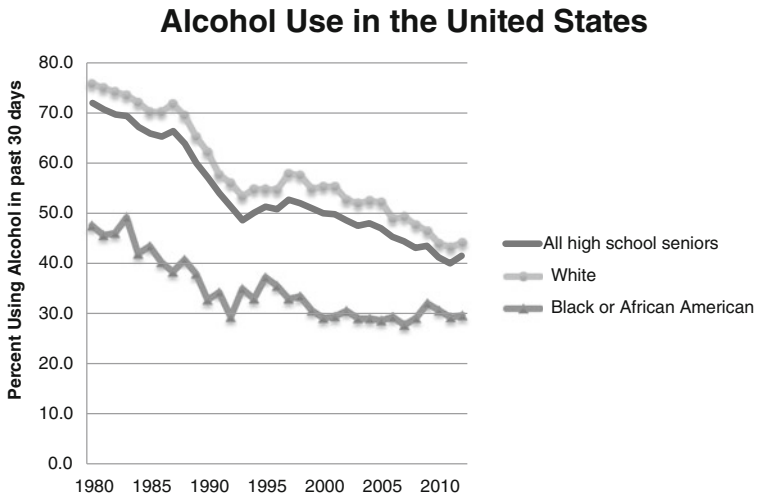


Fig. 9 Alcohol use in US high schools

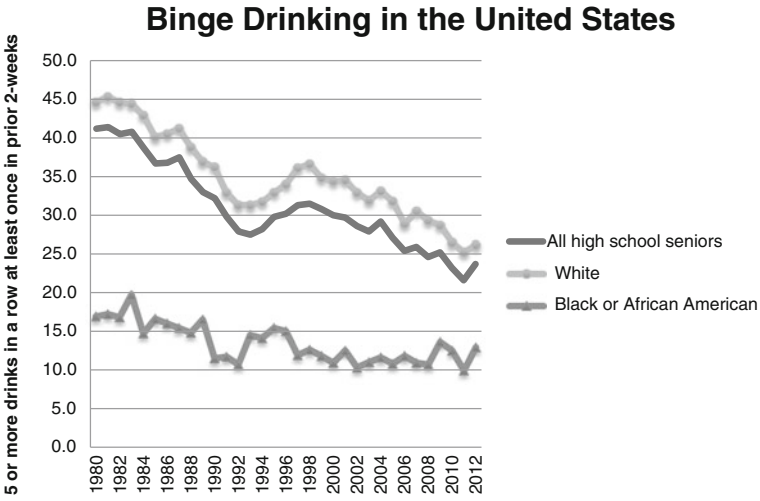


Fig. 10 Binge drinking in US high schools

two comparable datasets, the National Longitudinal Survey of Youth (US), 1997 and 2002.[42] Data for Canada comes from the similarly titled National Longitudinal Survey of Children and Youth, 2002 wave.[43] Here are the regression coefficients for optimism among the Québécois and African-Americans, holding constant income and the nurturing qualities of the mother. All children studied live with their mothers. Again, the Québécois are those who live in the province, speak French, and are white. The dependent variable in both the papers is some version of optimistic (Tables 4 and 5).[44]

African-American adults continue to be relatively more optimistic as well.[45] One study attributes the optimism, as well as the lower levels of suicide, to values consistent with Black culture: to report that God is responsible for life and to hold communitarian rather than individualistic values.[46]

1.3.1 The Common Problem and the Two Approaches: A reprise

Both the governments have a common problem under study here. This issue concerns what should be done with increasing cohabitation among a minority

[42]Ohio State University, Center for Human Resource Research, National Longitudinal Survey of Youth, 1997 (NLSY97), for which data is available at the Institute for Social Policy Research. http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/3959/detail. *N* = 1984. Unfortunately, the comparable data is not available for later years. The (US) General Social Survey in 2006 asked "I am always optimistic about my future."

[43]National Longitudinal Survey of Children and Youth. URL: http://www.statcan.ca/cgi-bin/imdb/p2SV.pl?Function=getSurvey&SDDS=4450&lang=en&db=imdb&dbg=f&adm=8&dis=2. The National Longitudinal Survey of Children and Youth (NLSCY) is a long-term study of Canadian children that follows their development and well-being from birth to early adulthood. The NLSCY began in 1994 and is jointly conducted by Statistics Canada and Human Resources Development Canada. Please note that the publicly available data used here is a synthetic dataset only. (N = 538). More recently, there has been a single question "How do you feel about your life as a whole right now?" asked in 2008, 2009, and 2010 that immediately precedes the Community Health Survey. This is available at select centers in Canada.

[44]The precise question in the NLSY97 is R0624200 (NLSY97). "I'm always optimistic about my future." (AGREE/DISAGREE). Answers ranged from 1 (Strongly Disagree) to 4 (Strongly Agree). The precise question in the NLSCY is EAMCcQ03 (NLSCY): "The next five years look good to me." Answers again ranged from 1 (Strongly Disagree) to 4 (Strongly Agree). The question is available at Questionnaire 10–19 Year Olds, Cycle 5, at 25 available at www.utoronto.ca/datapub/codebooks/cstdli/nlsc/synthetic/cycle5/nlysclc5-cbk-10-19-mas.pdf.

[45]See, e.g., Hope Yen and Jennifer Agresta, African Americans & Hispanics More Optimistic About Their Economic Future Than Whites, Poll Says, Huffington Post, March 27, 2015, available at http://www.huffingtonpost.com/2013/08/01/african-americans-hispanics-optimistic-economic-future-_n_3690867.html; Breanna Edwards, Survey: African Americans Still Optimistic Despite Racism, The Root, April 4, 2014, available at http://www.theroot.com/articles/culture/2014/04/survey_african_americans_still_optimistic_despite_racism.html.

[46]Rheeda L. Walker and Kelci C. Flowers, Effects of Race and Precipitating Event on Suicide versus Nonsuicide Death Classification in a College Sample, 41 Suicide and Life-Threatening Behavior 12 (2011).

Table 4 Optimism among Québécois adolescents

Model		Unstandardized coefficients		Standardized coefficients	*t*	Sig.
		B	Std. error	Beta		
1	(Constant)	2.274	0.142		15.993	0.000
	Québécois	−0.300	0.090	−0.136	−3.350	0.001
	Estimated total household income/poverty ratio	0.000	0.000	0.090	2.205	0.028
	Nurturing	0.048	0.006	0.316	7.768	0.000

R^2 (adjusted) for equation = 0.115, F = 24.368, sig. = 0.000 (optimistic.spv)

Table 5 Optimism among African-American adolescents

Model	Unstandardized coefficients		Standardized coefficients	*t*	Sig.
	B	Std. error	Beta		
(Constant)	2.773	0.065	0	42.347	0.000
Black	0.030	0.028	0.018	1.065	0.287
Household income to poverty ratio Percentage	0.020	0.005	0.069	4.186	0.000
Nurturing (residential mother supports child)	0.000	0.000	0.031	1.946	0.052

R^2 (adjusted) = 0.005, F = 7.518, sig. = 0.000

population, many of whom are poor. The Canadian solution, as I have stated it, is to stop privileging marriage, that is, to provide equal benefits to all who cohabit, thus recognizing de facto unions.[47] The solution in the United States first involves leaving solutions up to the individual states. Second, it privileges marriage. Examples include the federal marriage initiative,[48] which stresses marriage education, and the Defense of Marriage Act,[49] in which the federal government refuses to recognize same-sex marriage and through which states are freed from the usual obligation of honoring other states' marriages.

[47]C-23, RSC 4 (2d Supp.), SC 2000, c. 12.

[48]Section 101 of Pub. L. 104–193, 1996. The fruits of the legislation can be found at Office of Family Assistance, http://www.acf.hhs.gov/programs/ofa/programs/healthy-marriage.

[49]The Defense of Marriage Act (DOMA), Public Law No. 104–199, 110 Stat. 2419, 1996. This was invalidated I part United States v. Windsor, 570 US—, 133 S. Ct. 2675 (2013). The remainder is under challenge as Oberfell v. Hodges, 135 S. Ct. 1034 (2015).

1.3.2 Why Might the Outcomes Be So Different? Religion in Quebec: Policies and Reactions

In the 1990s, the Roman Catholic religious hierarchies in Québec that had performed most educational and health care services in the province ceded authority over them to the provincial government.[50] (For schools, this was a gradual process that did not conclude until 2006.)[51] Québec is now the only Canadian province with no church-run (parochial) elementary schools,[52] and religion may not be taught in or after school classes.

Beginning in the 1960s, church attendance in Quebec declined from the highest to the lowest rates in Canada.[53] While this may be for a number of reasons, some academics speculate that while Church reforms following Vatican II empowered the laity,[54] this movement did not relax unpopular stances toward birth control, abortion, and women's place in the Church.[55] Quebec now resembles some northern

[50]Alain Bélanger and Pierre Turcotte-Milan, "L'influence des caractéristiques sociodémographiques sur le début de la vie conjugale des Québécoises," 28 Cahiers québécois de démographie, 173 (1999). See also Lawrence Anderson, Federalism and Secessionalism: Institutional Influences on Nationalist Politics in Québec, Nationalism and Ethnic Politics, 13: 187–211 (2007).

[51]Comité sur les affaires religieuses, Secular Schools in Québec: A Necessary Change in Institutional Culture (October 2006), available online at http://www.mels.gouv.qc.ca/sections/publications/publications/BSM/Aff_religieuses/Avis_LaiciteScolaire_a.pdf (last visited April 22, 2011). See also Nugent, Demography, National Myths, and Political Origins: Perceiving Official Multiculturalism in Quebec, 38 *Canadian Ethnic Studies* 21 (2006).

[52]It is difficult to prove this negative, but see, for example, the Ontario Catholic School Board, which governs its system of (publicly funded) parochial elementary schools. http://www.tcdsb.org/ (last visited April 22, 2011). The only Catholic schools listed on the Archdiocese of Quebec's website are secondary schools. The website of the archdiocese of Quebec lists its various functions at beta.ecdq.org, and contains no references to schools in either its youth or formation subpages. The change came through the Education Act of 1988, Bill 107, 188 c 84 § 36. For a critical Roman Catholic perspective, see Rory Leishman, The School War in Quebec, Catholic Insight http://catholicinsight.com/online/church/education/article_1033.shtml (last visited April 22, 2011) (originally published in the Catholic Pro-Life publication The Interim, August 2010. Quebec was able to accomplish secularization by receiving an exemption from the Canadian Constitution that protected the two systems (religious and secular) in the other provinces. See Comité, supra note 60. For a discussion of the exemption, see David Cameron and Jacqueline D. Krikorian, Recognizing Quebec in the Constitution of Canada: Using the Bilateral Constitutional Amendment Process, 58 University of Toronto Law Journal 389 (2008).

[53]Warren Clark, Pockets of Belief: Religious Attendance Patterns in Canada, Canadian Social Trends 2, 3 (Spring,2003), Statics Canada Catalogue No. 11-008, http://www.statcan.gc.ca/pub/11-008-x/2002004/article/6493-eng.pdf.

[54]David Seljak, Why the Quiet Revolution was "Quiet": The Catholic Church's Reaction to the Secularization of Nationalism in Quebec after 1960, CCHA, 62 *Historical Studies* 109 (1996).

[55]Michael W. Higgins, The Bishop-maker: Who is Canadian Cardinal Marc Oullet? Commonweal Magazine July 22, 2010, www.Commonwealmagazine.org/bishop-maker (last visited April 22, 2011)(provides an accessible and brief story of the Quebec transformation and a discussion of current attitudes toward abortion in Canada).

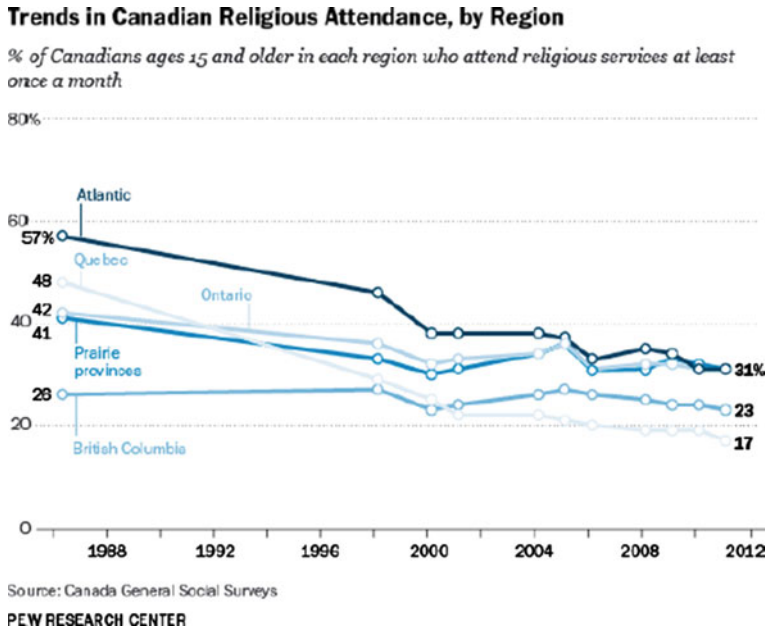


Fig. 11 Religious attendance, Canada

European nations in terms of attendance and importance given to religion.[56] In 1986, nearly half (48 %) of Quebec residents said they attended religious services at least once a month. By 2011, about one-in-six Quebecers (17 %) reported attending religious services at least once a month, a drop of 31 points (or 70 %), according to a Pew Research Center poll (Fig. 11).[57]

While the identity of the Québécois has been consistently and insistently oriented around the French language, the association with the Catholic Church has disappeared. Alain Bélanger attributes this to a rejection of what Michel Brunet called "les trois dominantes de la pensée canadienne-française: l'agriculturisme, le messianisme et l'anti-étatisme" [the three main components of French Canadian thought: agriculturalism, antistatism, and messianism].[58]

[56]Benoît Laplante, The Rise of Cohabitation in Quebec: Power of Religion and Power over Religion, 31 Canadian Journal of Sociology 1 (Winter, 2006).

[57]Pew Research Center, Canada's Changing Religious Landscape, at 11, available at http://www.pewforum.org/2013/06/27/canadas-changing-religious-landscape/.

[58]Alain Bélanger and Pierre Turcotte-Milan. 1999. "L'influence des caractéristiques sociodémographiques sur le début de la vie conjugale des Québécoises," *Cahiers québécois de démographie, 28: 173.* In English, see Claude Bélanger, http://faculty.marianopolis.edu/c.belanger/quebechistory/events/quiet.htm (last visited April 22, 2011).

1.4 A Recap of the Situation in the United States: Effect of Religiosity

The federal government in the United States provides much less generous social welfare support than does Canada.[59] In the United States, people who are married enjoy extensive legal and social protections.[60] People who are not must rely primarily on contract (or sometimes local domestic partner laws).[61] Unlike Canada, where the distinctions between married and common law couples are legally blurred (at least while the relationships last), the two categories remain quite distinct in the US. As we have seen, African-Americans, like the Québécois, do not take advantage of marriage to the same extent as the majority population. Most African-American children grow up in a family that, at least at some point, and sometimes from the beginning, is headed by a single mother. Andrew Billingsley and Barbara Morrison-Rodriguez argued that African-American communities turn to the church when they go through extending crisis for different types of support.[62] This may be particularly true for single mothers, often under stress.[63] Susan Sullivan found that mothers use religion to help build their children's self-esteem and give them "a sense of self-efficacy stemming from religious beliefs and prayer."[64]

One obvious difference is that while the United States is an outlier in terms of religiosity among "first world" nations, African-Americans, as a group, are far more religious than most.[65] This provides a contrast to the failing religiosity of the Québécois (Fig. 12).

---

[59]While this may seem obvious, support for this point may be found at Dennis Raphael and Toba Bryant, The Welfare State as a Determinant of Women's Health: Support for Women's Quality of Life in Canada and Four Comparison Nations, 68 Health Policy 64, 64, 68 & Table 9 (labor market public spending) (2004).

[60]One list of what these are appears in the Vermont case that eventually resulted in the enactment of civil union legislation, Baker v. State, 744 A.2d 864, 884–84 (Vt. 1999).

[61]See, e.g., Patricia A. Cain, Imagine There's No Marriage, 16 *QLR* (Quinnipiac) 27 (1996).

[62]Andrew Billingsley and Barbara Morrison-Rodriguez, "The Black Family in the 21stCentury and the Church as an Action System: A Macro Perspective." 1 Journal of Human Behavior in the Social Environment 1:2-3, 31–47 (1998).

[63]Susan Crawford Sullivan "The Work-Faith Connection for Low-Income Mothers: A Research Note." 67(1): Sociology of Religion 99, 106 (2006).

[64]Susan C Sullivan, "Unaccompanied Children in Churches: Low-Income Urban Single Mothers, Religion, and Parenting." 50(2) Review of Religious Research 157, 170 (2008).

[65]See, e.g., Rheeda L. Walker, David Alabi, Jessica Roberts, and Ezemenari M Obasi, Ethnic Group Differences in Reasons for Living and the Moderating Role of Cultural Worldview, 16 Cultural Diversity and Ethnic Minority Psychology 372, 373 (2010).

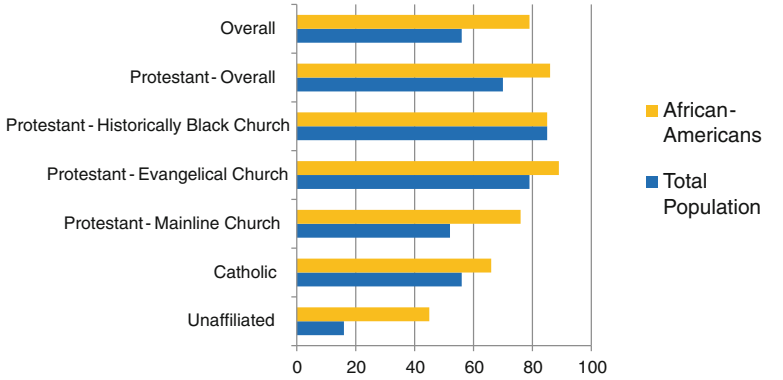


Fig. 12 Religiosity of African-Americans [A Religious Portrait of African Americans, Pew Forum on Religion & Public Life, Jan. 30, 2009, http://pewforum.org/A-Religious-Portrait-of-African Americans.aspx. (last visited April 22, 2011)].

2 Conclusion: The Importance of Support

Families need community support in order to function well, particularly when we consider children's well-being.[66] Typically, this support comes from communities through the legal status of marriage and adoption. However, while it is not optimal, cohabitation may suffice for children's well-being (though it will not be stable), but only if the parents (in most cases, the mothers) have some sort of other, external support. In this chapter, I have tried to demonstrate that religion appears to be a mediating communitarian factor for African-Americans but not for the Québécois. This difference may explain the better psychological, health, and mortality outcomes.

I would like to interject a few words of caution, however. The empirical comparison drawn here is only for a few outcome variables, and, at least for my own work on optimism, uses the simplest possible model. In fact the Canadian data is a "synthetic" dataset rather than the complete one. The consistent findings with depression, suicide, and tobacco use should support this interpretation, however.

References

Beaupré, Pascale, P. Turcotte and F. Goldscheider. 2005. Stablité des couples en union libre avec enfants: Le Québec se distingue-t-il des autres provinces canadiennes?. Powerpoint May Slide

[66]The need for support, or social capital, is one of the pervasive themes of Margaret Brinig, Family, Law and Community, supra note 38, especially Chapters 1 and 2. For a book explaining its role in social mobility, see Robert Putnam, Our Kids: The American Dream in Crisis (Simon and Schuster 2015).

6, available (in French), Quebec Inter-University Center for Social Statistics, poster presented at Population of Association of American Meeting, Boston, Massachusetts, April 1, 2004. www.ciqss.umontreal.ca/Docs/Conference/Mai2005/Turcotte_Mai2005.pdf. Accessed 21 April 2011.

Légaré, Gilles, M. Gagné and D. St-Laurent. 2014. *La mortalité par suicide au Québec: tendances et données récentes*—1981 à 2011. Institute National de Santé Public du Quebec. http://www.inspq.qc.ca/publications/notice.asp?E=p&NumPublication=890. Accessed 21 April 2011. at 5 & Fig. 3. (The adolescent rate has the open dashes.) While suicide rates have generally declined, depression has not. Cheung & Dewa, note 33, at 82 and Table 1, 84 & Table 3.

Marcil-Gratton, Nicole and Celine de Bourdais. 1999. *Custody, access and child support: Findings from the national longitudinal survey of children and youth*. http://www.justice.gc.ca/eng/pi/fcy-fea/lib-bib/rep-rap/1999/anlsc-elnej/index.html.

Walker, Rheeda L., and Kelci C. Flowers. 2011. Effects of Race and Precipitating Event on Suicide Versus NonSuicide Death Classification in a College Sample. *Suicide and Life-Threatening Behavior* 41: 12.

Marquis de Condorcet and the Two-dimensional Jury Model

Manfred J. Holler

1 The Condorcet Program

On March 27, 1794, after hiding several months in Paris, Marquis de Condorcet was found out and arrested, perhaps betrayed, and imprisoned at Bourg-la-Reine. The next morning he was found dead, either by poison or from exhaustion, a few months before the Jacobin terror ended as Robespierre was executed himself. Condorcet was under proscription as Girondin, and considered a challenge to Robespierre and his regime. This is a one bare-bone story of the end of the life of a genius. In fact, alternative dates and stories of his death are published. However, in this paper we are interested in his earlier work which was motivated by the vision and urge to extend the domain of reason to social and political affairs. In his 1785 *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix* (*Essay on the Application of Analysis to the Probability of Majority Decisions*), he discussed (a) the efficiency of a jury decision to approximate the truth through majority voting, now called Condorcet's jury theorem, (b) a paradox in the aggregation of preferences, the Voting Paradox also called Condorcet paradox, and (c) a method of voting which consists of pairwise voting decisions between all candidates (or alternatives) in an election. This method results

I would like to thank Hannu Nurmi for very helpful comments.

Parts of this paper, especially the model discussed in Sects. 3–5, are borrowed from Holler and Napel (2007) and Holler (2010).

M.J. Holler (✉)
Center of Conflict Resolution (CCR), University of Hamburg,
Von-Melle-Park 5, 20146 Hamburg, Germany
e-mail: manfred.holler@wiso.uni-hamburg.de

M.J. Holler
Public Choice Research Centre (PCRC), University of Turku, Turku, Finland

A. Marciano and G.B. Ramello (eds.), *Law and Economics in Europe and the U.S.*,
The European Heritage in Economics and the Social Sciences,
DOI 10.1007/978-3-319-47471-7_9

either in selecting the so-called Condorcet winner if there is a candidate who wins a majority in each pairwise comparison, or in a cycle, i.e., the Condorcet paradox. All three concepts of voting are relevant in what follows. They define the Condorcet program.

In Sect. 2 details of Condorcet's jury theorem are given and discussed with respect to applications in the sociopolitico context in search of the *volonté générale.* If there is a *volonté générale*, should a majority be able to find it? Or, is what the majority finds the *volonté générale*? Specifying the idea behind these questions, Sects. 3 and 4 present and analyze a model that explicitly combines the idea of a jury with the aggregation of preferences. What looks like a toy model could be seen as a close approximation of an investment arbitration court proposed by the Transatlantic Trade and Investment Partnership (TTIP), in the course of an Investor-State Dispute Settlement (ISDS). Section 5 discusses the model, its assumption, and equilibria, and offers some more general results, pointing out a second-mover advantage in its decision structure. Section 6 summarizes the discussion relating Arrow's work with Condorcet's program. A remark to TTIP concludes this section.

In Condorcet we find both: the belief in a generalizable truth, which is a heritage of his deep involvement in enlightenment, and the liberal idea which acknowledges the individual as the building block of society and therefore also of social decision-making. Emma Rothschild (2001, p. 197) calls him "an interstitial figure…. He belongs to neither side, entirely, in the philosophical dichotomy of uniformity versus diversity, or of universal connectedness versus the endlessness of conflict." There is room for specification and interpretation. The following sections make use of this potential.

2 Jury Theorem and *Volonté Générale*

Condorcet's jury theorem says that (i) any jury of odd number of jurors is more likely to select the correct alternative than any single juror; and (ii) this likelihood becomes a certainty as the size of the jury tends to infinity. The theorem holds if (a) the jury N decides between two alternatives by voting under simple majority rule; (b) each juror i has a probability $p_i > ½$ to be correct; (c) $p = p_i$ for all i in N; and (d) each juror i decides independently. (See Boland 1989; Grofman et al. 1983.)[1] A probability $p_i > ½$ defines an expert. Unfortunately, these four assumptions hardly ever (or, most likely, never) hold in reality and therefore increasing the number of jury members is not always a reliable instrument to come closer to the truth. As demonstrated by Kaniovski and Zaigraev (2011), the optimal

[1]For Condorcet's text on the jury theorem, see Condorcet (1785, pp. 119–136) or, translated into English, Sect. 11 in Condorcet (1989), specifically p. 107. For a modern theoretical analysis and application, see, e.g., Kaniovski (2008, 2010) and Kaniovski and Zaigraev (2011).

jury size may imply a single juror if simple majority rule applies, all jurors are equally competent, but competence is low, and correlation between the jurors is high.[2]

The jury theorem is well known among scholars of Law and Economics and references are ubiquitous. What is less known is that Condorcet tried to extend his probability approach to the aggregation of preferences, and failed.[3] However, this experiment left us with the Voting Paradox, Condorcet's second outstanding contribution. It did not only inspire Arrow (1963 [1951]) to write his *Social Choice and Individual Values*, but also triggered earlier work that presented the two-dimensional jury model discussed below (Holler 1980, 1982, 2010; Holler and Napel 2007).

In principle, aggregation of preferences is not about finding some truth, but to summarize the evaluation of feasible or available alternatives, possibly social states. Therefore, the jury theorem does not apply and its probability calculation seems, at least at the first glance, to be vacuous. Black (1963, p. 163) concludes "whether there be much or little to be said in favour of a theory of juries" that refers to probability calculation, "there seems to be nothing in favour of a theory of elections that adopts this approach." He adds "…the phrase 'the probability of the correctness of a voter's opinion' seems to be without definite meaning." However, Arrow (1963 [1951], p. 85) gives a somewhat surprising interpretation of Rousseau's *volonté générale* and voting: "Voting, from this point of view, is not a device whereby each individual expresses his personal interests, but rather where each individual gives his opinion of the general will." And he concludes, that this "model has much in common with the statistical problem of pooling the opinion of a group of experts to arrive at a best judgement…"[4]

This could be interpreted as a justification of using juries of experts to choose the winner in competitions in the fields of arts and sports. However, legal judgements are not always about finding or defining the truth. Often they are about what is good or bad, or what should be done and what should be omitted, and there are degrees of the goodness and badness of the alternatives to be judged. Almost every member of the corresponding society is considered an expert in this field. It cannot be denied that some justification for this can be found in the argument that relates Rousseau's *volonté générale* to voting.

Condorcet (1785, p. xxv) points out that there are objects for which an immediate personal interest will light the spirit: the maintenance of safety, freedom, and property. On such issues a direct democratic vote can be taken by a large assemble, i.e., the nation, without damage. On other issues, decisions taken by unenlightened

[2]Ladha (1993) introduces symmetrically dependent votes and demonstrates that Condorcet's main result holds: a majority of voters is more likely than any single voter to choose the better of two alternatives. Berg (1993) discusses correlation by assuming that a juror's competence depends on preceding votes.

[3]See Black (1963, pp. 64ff.) for this judgement and the arguments.

[4]Later, the relationship of voting and Rousseau's common will was excessively discussed. See Grofman and Feld (1988) and the literature given in this article.

large numbers of voters can be dangerous—"unless the subjects under debate are very simple" (Condorcet 1989, p. 107). Voters have to be experts otherwise large numbers do not approximate truth, but disaster. Condorcet's argumentation indicates that, in principle, majority voting can be applied to issues that refer to value judgements as long as the corresponding values are widely shared in an enlightened nation. However, for issues of some complexity the condition of expertise, assumed by the jury theorem, puts constraints and limitations on the democratic process and the application of Jakob Bernoulli's principle of large numbers. Condorcet expects that the larger the number of decision makers, the lower will be their competence, in general, and therefore collective decisions resulting in useful reforms of the administration or of law making, will become less likely, the larger the number of voters. (See Condorcet 1785, p. xxv.)

In addition, individual preferences on social values are likely to be correlated and the jury theorem in its proper form does not apply.[5] And still, social judgements on values presuppose either the existence of a scale of values, i.e., a social welfare function, or a mechanism that brings about an evaluation scale or the choice of a particular alternative. A jury is such a mechanism. On the one hand, juries are used to decide on rank orders in competitions. On the other hand, they decide on guilty and nonguilty, or select from a bundle of alternatives the duties that a convict has to accomplish. In general, however, they do not guarantee that the decisions satisfy an Arrow-type social welfare function if there are more than two alternatives. Moreover, such a welfare function may not exist. Arrow (1963 [1951]) demonstrated that there is no social welfare function, i.e., a "process or rule" that maps the set of individual preferences profiles into the set of social preference orderings, both defined on the same sets of alternatives, such that it satisfies two well-known axioms and five "reasonable" conditions. The conditions are: (i) "unrestricted domain" which says that none of the possible preference profiles on the given set of alternatives should be excluded; (ii) "monotonicity" which refers to Paretian efficiency ["Since we are trying to describe social welfare and not some sort of illfare, we must assume that the social welfare function is such that the social ordering responds positively to alterations in individual vales" (Arrow 1963 [1951], p. 24).], (iii) "independence of irrelevant alternatives", (iv) "citizen sovereignty" which in Arrow's words implies that the social welfare function is not "imposed", i.e., it derives from individual preferences; and (v) "nondictatorship". Condition (v) says that there is no decision maker *i* so that the social preferences are identical with the preferences of *i*, irrespective of what the preferences of the other members of the society are, whether and how they change.

Of course, if the domain is restricted, e.g., by the implications of preferences which are single-peaked, then the aggregation could result in a social welfare function that satisfies (ii) to (v). Perhaps it is sufficient to assume that the members of the society are enlightened. It seems that was Condorcet's creed. However, he

[5]Kaniovski (2010) demonstrates that an enlargement with positive correlation can be detrimental up to a certain size, beyond which it becomes beneficial.

also admits that this is an ideal, never been observed so far in history. (See Condorcet 1785, p. xxv.) Still, he believed in the power of education and the progress of society.[6]

Arrow postulates that the social welfare function should satisfy the very same axioms that define individual preference orderings: "connectivity" and "transitivity" where "connectivity" implies both "completeness" and "reflexivity" which are standard for the definition of an individual preference ordering. To restate, his theorem says that there is no social welfare function that satisfies these properties and the five conditions listed, i.e., that allows to socially rank each pair of alternative no matter what the individual preferences say. The model discussed in the following sections illustrates the problem of a "missing social welfare function." It analyzes a situation and a procedure of preference aggregation that is however conclusive inasmuch as it selects a winning alternative for almost all preference profiles of the jury.

3 The Two-dimensional Jury Model

The decision situation and the procedure of preference aggregation are specified by a two-dimensional model which is based on the so-called Holler–Steunenberg model discussed in McNutt (2002, pp. 282ff) and applied to European decision-making in Holler and Napel (2007). It has its roots in Holler (1994) and Steunenberg (1994). It assumes a sequential structure of decision-making that is quite similar to the ultimatum game. There is a proposer and a responder. However, the game below endogenizes the judgements that characterize the empirical results of the ultimatum game that indicate a deviation of the subgame perfect equilibrium (when utilities are assumed to be linear in money). What is attributed to concerns of justice and envy in the interpretation of the ultimatum game is institutionalized by a jury. In fact, this is perhaps the most important function of juries: to institutionalize judgements that are meant to be based on justice (or truth).

The model combines the idea of winning a maximum of votes in a voting game (i.e., the jury) with utility maximization that derives from the winning proposition. The model assumes a first mover A, the agent of the plaintiff, and a second mover D, the counsel of the defendant. In what follows we call A the "plaintiff" or proposer, and D the "defendant" or responder. Typically, A and D are agents of parties that have conflicting interests. They face a jury that consists of three voters: $J = \{1, 2, 3\}$. Each voter is characterized by strict ordering of preferences on the set of alternatives and is given by $\Omega = \{u, v, w\}$. Its elements describe the possible outcome selected by a simple majority of the jury members (i.e., the voters), subject to the alternatives presented by A and D.

[6]See Faccarello (2016) and Behnke (2011). However, Rothschild (2001, p. 198) observes that "Condorcet's political and philosophical opinions changed substantially in the course of his public life…He became considerably more skeptical, in particular…about the prospects for education in enlightenment…".

Table 1 represents the preference profile of the jury members. Voter 1 prefers u to v and v to w, and so on. Figure 1 demonstrates that the preferences of the voters are not single-peaked, i.e., the preference profile is intransitive while individual preferences are assumed to be transitive.[7] Note that Table 1 represents a selection of jury members with a maximum of diversity in their preferences. As observed by Rothschild (2001, p. 198), Condorcet "was preoccupied, in the first place, with individual diversity." Pairwise comparison of alternatives implies cyclical majorities as there is no Condorcet winner if voters vote sincerely, i.e., if they vote in accordance with their preference orderings expressed in Table 1. As a consequence, if agent A proposes alternative $x \in \Omega$ there is always an alternative $x' \in \Omega$ that is preferred to k by a majority of jury members if presented by agent D.

If A is interested in winning a majority and thus to win the case, and D has the same target, then its intentions will be frustrated whatever alternative A proposes. However, in general, legal cases are not only about winning, but also about outcomes. The clients of A and D have preferences with respect to the elements of Ω and their agents A and D have to take these preferences into account. We assume that A and D represent the preferences $w \succ u \succ v$ and $u \succ v \succ w$, respectively. (Here, symbol $\succ$ represents the binary relationship "better".) This defines the first dimension of the agents' preferences.

The second dimension of their preferences is indeed defined by "winning," "losing," and "compromise." We assign the numbers 1, 0, and ½ to these events. Given a particular outcome $x \in \Omega$, both A and D prefer event 1 to event ½ and event ½ to event 0. Often, in a legal case, the losing party has to pay fees to the court and cover the legal expenditures of the winning party. Therefore, winning the case can be beneficiary per se.

More generally, we can write the preferences of the two agents A and D in the form of a utility function $u_i = u_i(m, p)$, $i = A, D$. Here, $m \in M = \{0, ½, 1\}$ expresses the probabilities of winning of a majority of votes in the jury which is assumed to be ½ in the case of the indifference of the decision makers or in the case of nondecisiveness (ties) in the voting body. Or, it signals that both parties agree on a specific alternative. This alternative can be understood as a compromise with the consequence that the case will be closed and no vote is taken. Thus, there will be no loser and no winner. The variable p is defined by $p \in P = \{u, v, w\}$. Here, P describes the *discrete* set of alternatives that the agents can choose. We assume that this set is identical to the set of alternatives that can be submitted to a vote. Thus the sets P and Ω are identical.[8]

We further assume that agent A knows the preferences of D, and agent D knows the preferences of A, and both know the preferences of the jury members as shown

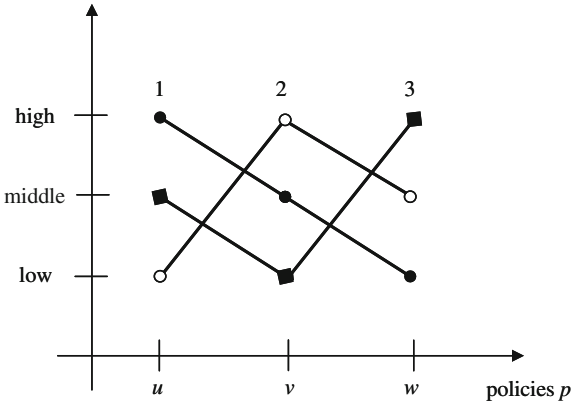
[7]There is no ordering of u, v, and w such that the preferences of all voters are single-peaked. Thus, the preferences are nonsingle-peaked (see Black 1948).

[8]The variable m represents the standard vote maximizing objective that public choice theory assumes for political agents, while p is a close relative to the utility maximization suggested in Wittman (1973) that becomes relevant if the incumbent (i.e. the proposer) faces cyclical majorities and thus cannot win an election.

Table 1 Preference profile of the jury members

Ranking	Voter 1	Voter 2	Voter 3
High	*u*	*v*	*w*
Middle	*v*	*w*	*u*
Low	*w*	*u*	*v*

Fig. 1 Nonsingle-peaked preferences



in Table 1. The assumption that *A* and *D* know the preferences of the other party is perhaps not far away from most real-world settings. However, knowing the preferences of the jury members seems to be more daring. However, given these assumptions, the game model that is discussed in the following is characterized by complete (and perfect) information. We now derive the optimal choices of *A* and *D* in this game. This problem is "solved" for a subgame perfect equilibrium by backward induction. Agents *A* puts himself into the "shoes" of *D* and ask how will *D* react if *A* presents *u*, *v*, or *w*, alternatively. The choices of *A* are represented by *u**, *v**, and *w** in Figs. 2 and 3. What are the best replies of *D*, given the choices by *u**, *v**, and *w**?

4 The Optimal Choices

The potential best reply set of agent *D*, illustrated in Fig. 2, shows the outcomes derived from the choices of *A* and *D* for the given preferences.[9] If *A* chooses *w** and

[9]These preferences result from applying the dominance relation, but do not consider trade-offs between *m* and *p*. For example, in *g*'s perspective *u* dominates *v* as a response to *a*'s choice of *u* (denoted by *u** in Fig. 2). However, winning for sure ($m = 1$) with w, that is ceteris paribus the least desirable policy for *D*, may potentially be preferred to responding with *u* resulting in outcome *u* (indicated in bold in Fig. 2) but only $m = 1/2$.

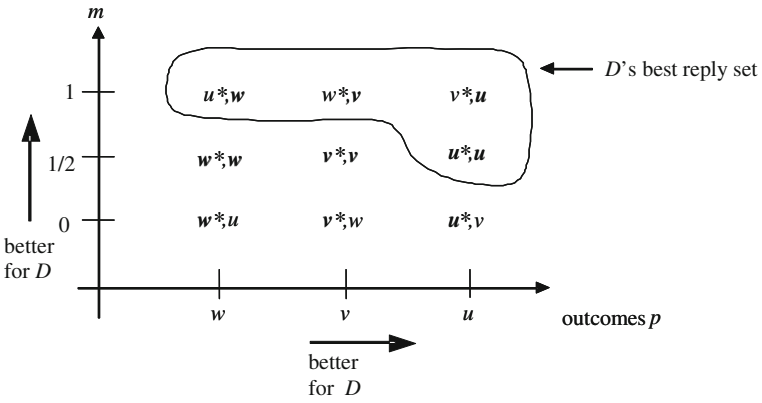


Fig. 2 Best reply set of agent D

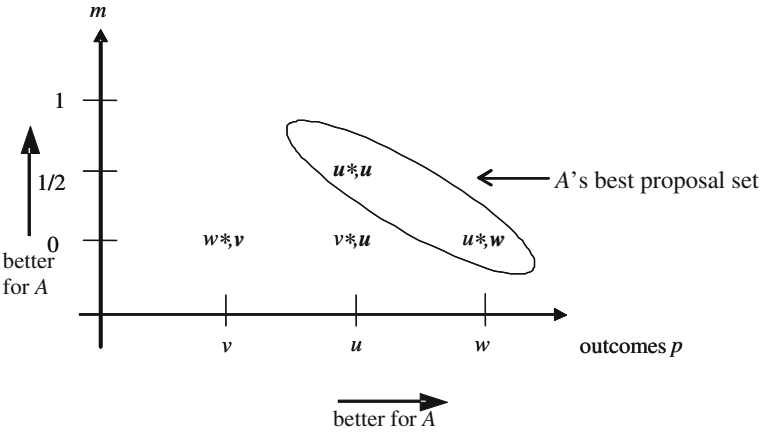


Fig. 3 The best proposal set of agent A

D selects v, voters 1 and 2 will vote for v and 3 will vote for w (see Fig. 1). Thus D will win a majority of votes $(m = 1)$ and v will be the outcome.

The ranking of D on the pairs (m, p) is illustrated in Fig. 2. For example, D prefers the outcome $(1, \mathrm{v})$ to $(1, w)$ which results from the choices represented by (u^*, w). However, D prefers $(1, u)$, which results from the choices (v^*, u) to $(1, v)$. Agent D's potential best reply set shows the outcomes which derive from the choices of A and D for the given preferences. If A chooses w^* and D selects v, voters 1 and 2 will vote for v and 3 will vote for w (see Fig. 1). Thus D will win A majority of votes $(m = 1)$ and v will be the policy outcome. The ranking of D on the pairs (m, p) is illustrated in Fig. 2. For example, D prefers the outcome $(1, \mathrm{v})$ to

$(1, w)$. The latter results from the choices (u^*, w). However, D prefers $(1, u)$, which results from the choices (v^*, u), to the outcome $(1, v)$.

Given $m = 1$, Fig. 2 reflects the Condorcet Paradox: D will win with certainty and no $x \in \Omega$, or, equivalently, $p \in P$ exists which can prevent D from winning. The pair (u^*, u) says that both A and D select policy u and thus there is a 1 in 2 chance of each of them winning the election.

Obviously, seen from the perspective of agent A, there are elements in the potential best reply set of D that are dominated by another element in this set. Figure 3 illustrates A's evaluation of the elements contained in D's potential best reply set. Given A's preferences, (w^*, v) and (v^*, u) are clearly dominated by (u^*, u) and (u^*, w). Thus, we can conclude that A will propose the alternative u^*. Whether D accommodates and proposes an identical policy or whether it selects w to defeat the proposed policy u^*, is a question of D's preferences on (u^*, u) and (u^*, w). If we abstract from the case that D is indifferent as regards these two alternatives, then the outcome of the two-dimensional jury game is uniquely determined and corresponds to a subgame perfect equilibrium.

More generally, every finite sequential-move game of perfect information has a *unique* subgame perfect equilibrium if all players have strict preference orderings over the possible outcomes. This follows by backward induction.

Note that the social preferences, i.e., voting outcomes, are not cyclical although we dropped the assumption of one-dimensional single-peaked preferences. Note further that the voting outcome could be u irrespective of whether A or D is winning the election. Thus we conclude that there is a chance for a rather stable arrangement despite the fact that voter preferences are nonsingle-peaked.[10] The platform u can function as a substitute for the median position which is not defined for cyclical preferences. This implication of the above model is quite different from the standard result in the case of nonsingle-peaked voter preferences which suggests that the winning outcome will strictly depend on the agenda in pairwise voting. For instance, given the preferences of the voters in Table 1 such that no alternative represents a majority of votes, w will be the outcome if u and v are submitted to voting in the first round and the winner, u, competes with w in the second round—just to get defeated by w. The proposer–responder structure of the model works similar to defining an agenda.

Holler (1982) analyzes all 36 cases that result from combining the possible preferences of a first mover A and a second mover D, if the preferences of the two candidates have the structure of any of the three preference orders given in Fig. 1. Each best proposal set of the corresponding proposer–responder game contains two undominated alternatives. One of these alternatives is characterized by a pair of identical propositions. This implies that there is a chance that the result will be the same, irrespective of the agent who wins a majority of votes. In the case discussed

[10]Laver and Shepsle (1990) analyze a two-dimensional voting model with discrete alternatives, rather similar to the model presented here, however with single-peakedness in each dimension. They show that, in general, the core is not empty—therefore stable outcomes are likely.

above, this of course presupposes that both agents prefer (1/2,*u*) to winning a majority "with certainty" but having to propose something less preferred than *u*. The latter possibility characterizes the second undominated alternative in the best proposal set.

From the analysis of 36 cases in Holler (1982) we can conclude:

(i) There is a second-mover advantage in the above game: being the first to present a proposal can never be preferred to being the second. If the proposal of *A* is acceptable to *D*, because it ranks high in *D*'s preference order, then the latter can select an identical proposition, thereby gaining a 50 % chance of winning the election. If the proposal of *A* is not acceptable to *D*, because it ranks low in *D*'s preference order, *D* can present a different proposal and win a majority of votes.

(ii) However, there are combinations preference profiles for jury members and agent's preferences on $P = \{u,v,w\}$ such that the outcome of *A* presenting a proposal first and *D* second are identical to the outcomes of *A* presenting a proposal second and *D* presenting a proposal first. That is, the second-mover advantage is "weak".

5 Discussion of the Model

In this section, we will discuss our results with reference to two standard models. First, we relate them to Don Saari's observation that a majority cycle profile is not neutral when matched with other preferences. Then we ask the question whether our results are different from the standard observation that the agenda is decisive for the selection of the winner, given cyclical majorities. Following Saari (1995)[11], we now combine our proposer–responder model with a jury that is distinguished by a preference profile that is in a way complementary with the profile in Table 1. It consists of the "other" three preference orderings that can be formed out of three alternatives. (There are n! = 6 different orderings that can be formed out of n elements.) Not surprisingly, the preference profile in Table 2 implies cyclical majorities as well.[12]

Now let us see the (optimal) choices of *A* and *D* facing the jury represented by Table 2. The best reply set of D is illustrated in Fig. 4. Again it illustrates *D*'s best replies on alternative propositions possibly brought forward by *A*. Starting from this result, we derive the best proposal set for *A*. The result is illustrated in Fig. 5.

[11]See Nurmi (2006, p. 131) for illustration and discussion.

[12]Saari's concept of a "ranking wheel" allows for identifying the preference profile that is characterized by a majority cycle, i.e., a voting paradox (see Saari 2011). For the case of three alternatives there are two (types of) profiles that imply a voting paradox.

Table 2 Preference profile with cyclical majorities

Ranking	Voter 1	Voter 2	Voter 3
High	*u*	*v*	*w*
Middle	*w*	*u*	*v*
Low	*v*	*w*	*u*

Figure 5 implies that *A* can initiate outcome (*u**,*u*) or outcome (*v**,*w*). If *A* prefers achieving its highest ranking alternative at the expense of losing the case through jury voting, to achieve its second ranking alternative and a chance of ½ to win the case, then A will propose *v*. Correspondingly, *D* will react with *w* and *w* (and *D*) will be the winner(s). If not, then *A* proposes *u* and the outcome will be alternative *u*, and *A* will win the case with probability ½. Obviously, given the jury's preference profile in Table 2, *A* decides what alternative will result. This implies a first-mover advantage for *A*. Note, in case that the jury's preference profile is given by Table 1, *D* decides whether *u* or *w* will be the outcome. This confirms that Condorcet paradox profiles are *not neutral*: The jury can have an impact on the final outcomes even if the preference profile of its members is intransitive and not conclusive applying majority voting or pairwise comparison à la Condorcet.

It is well known that in face of a Condorcet paradox the agenda decides on the outcome in case of pairwise voting. Given a preference profile as in Fig. 1 has a majority, alternative *u* will be the winner if *v* and *w* compete in a first round and *u* challenges the winner of this round. Similarly, *v* can be made winner if *u* and *w* compete in the first round. The above proposer–responder model endogenizes the agenda. It adds competition on selecting the alternatives. The competition results from the agents' interest in the resulting alternative and in the winning of a majority of votes. This model still allows a sequence of alternating alternatives to win that can be interpreted as a cycle, if *A* and *D* take turns, however, it does not exclude a stable result as suggested by (*u**,*u*) in the above specification.

6 Arrow Versus Condorcet and TTIP

In Arrow (1963, p. 1) we can read that in "a capitalist democracy there are essentially two methods by which social choices can be made: voting, typically used to make 'political' decisions, and the market mechanism, typically used to make 'economic' decisions." The procedure that we analyzed above does not give a social ranking of the alternatives, but indicates a possible choice. Thus, strictly speaking, it does not define a social welfare function but a social choice function. However, this outcome concurs with the result that we expect from applying voting procedures[13] and the function of making political decisions that Arrow assigns to

[13]There are however voting procedures that give a ranking of the alternatives that can be interpreted as a social welfare function (e.g., Borda count). Needless to say that such social welfare functions do not satisfy Arrow's axioms and conditions as stated above.

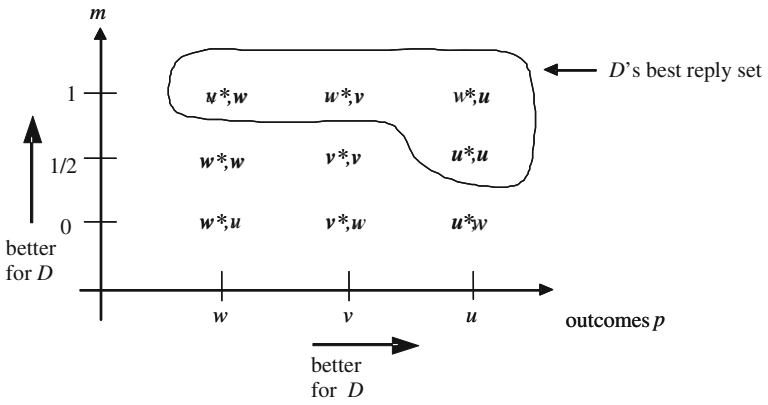


Fig. 4 Best reply set of agent *D*

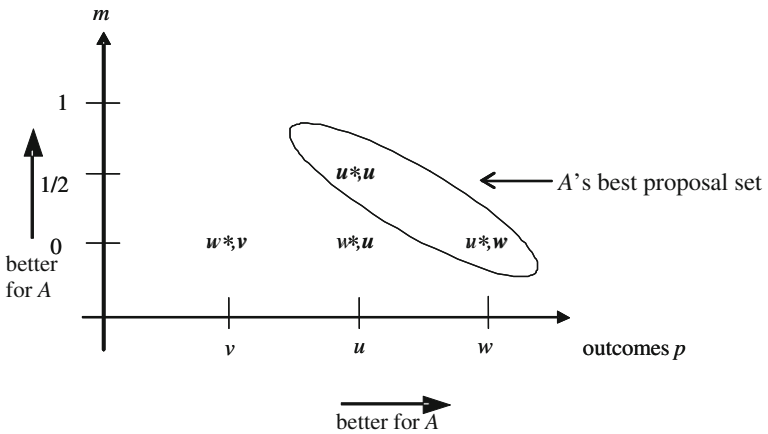


Fig. 5 The best proposal set of agent *A*

them. Voting procedures imply the counting and adding up of votes. This implies cardinality and interpersonal comparison, irrespective of whether "one person, one vote" applies or votes are weighted like, for instance, in the EU council of ministers. There is a fundamental tension between Arrow's project of a social welfare function that assumes ordinal preferences of the individuals and ordinality of the social ranking, and the cardinality of counting votes (i.e., counting votes)—and Arrow's assertion "…that interpersonal comparison of utilities has no meaning and, in fact, that there is no meaning relevant to welfare comparisons in the measurability of individual utility …If we cannot have measurable utility…, we cannot have interpersonal comparability of utilities fortiori" (Arrow 1963 [1951], p. 9).

Condorcet establishes that it is welfare, and not happiness, which is "a duty of justice" for governments. However, his concept of welfare is less demanding, from a theoretical point of view, than the one embedded in the Arrovian world. As summarized in Rothschild (2001, p. 201), for Condorcet welfare "consists in not being exposed to misery, to humiliation, to oppression. It is this welfare which governments owe to the people. It is necessary to happiness, and it may not be sufficient. But it is up to nature to do the rest."

Condorcet suggests collective decision-making on issues of public interest that specify issues of welfare, i.e., public institutions. But given the problems of aggregating preferences he came across in his studies, he suggests a careful design of the corresponding assembles: "We must therefore ensure that elective assemblies are formed in such a way that they only rarely produce a plurality which leads to a result of this kind" (Condorcet 1989, p. 76f). Most of his probability studies in the *Essai* (Condorcet 1785) can be interpreted as offering insights into such problems and ways to circumvent them.

Interestingly, Condorcet does not suggest the classical way of dealing with excessive heterogeneity in the population: streamlining the population by means of education. For instance, in ancient Greece the education of the youth was considered an effective instrument to infuse standards of moral behavior in accordance to the existing social norms. Education supplemented the political institutions. Further, since the social norms varied substantially among the various city-states—think about Sparta, on the one hand, and Athens, on the other—the education differed as well (for details, see Bitros and Karayiannis 2010). Rothschild (p. 199) writes: "In his last major published work, on principles of education, Condorcet again identified the diversity of opinions as a preeminent good. Public instruction should not extend to political and moral education, which would be contrary to the 'independence of opinions'. Children would still, perhaps, 'receive opinions' from their families. But these opinions would not then be 'the same for all citizens'. They would not have the character of a 'received truth,' or a 'universal belief'".

So we are back to the design of institutions, e.g., voting assembles, as a means to bring about "good social outcomes." A discussion of the two-dimensional jury model could be seen as a contribution to this project. In the introduction I suggested that this toy model could serve to discuss the installation of investment arbitration courts as proposed by the Transatlantic Trade and Investment Partnership (TTIP) in the course of an Investor-State Dispute Settlement (ISDS). In fact, ISDS provisions are already contained in a number of bilateral investment treaties and in international trade treaties, such as the North American Free Trade Agreement and the Trans-Pacific Partnership. They grant an investor the right to use dispute settlement proceedings against a foreign government. In general, disputes can be initiated by corporations and natural persons. In Wikipedia[14] we can read that "tribunals are composed of three arbitrators. As in most arbitrations, one is appointed by the investor, one by the state, and the third is usually chosen by agreement between the

[14]"Investor-state dispute settlement", version of February 9, 2016.

parties or their appointed arbitrators or selected by the appointing authority, depending on the procedural rules applicable to the dispute. If the parties do not agree whom to appoint, this power is assigned to executive officials usually at the World Bank, the International Bureau of the Permanent Court of Arbitration, or a private chamber of commerce." If we assume that tribunals decide by simple majority, then the above toy model is a close representation of reality. It seems natural to assume that the discontent investor is the plaintiff and the indicted state is in the role of the defendant. The above model suggests a weak second-mover advantage for the defendant. Do we want to see this advantage for an indicted state? Or should we arrange the preferences of the jury such that no second-mover advantage prevails? The model could help to discuss and perhaps even solve this problem—if it is solvable at all.

References

Arrow, K.J. 1963 [1951]. *Social choice and individual values*, 2nd ed. New York: Wiley.

Behnke, J. 2011. Condorcet und die 'soziale Mathematik'. Eine kurze Einführung in Leben und Werk. In: *Condorcet*, 1–48.

Berg, S. 1993. Condorcet's jury theorem, dependency among jurors. *Social Choice and Welfare* 24: 87–95.

Bitros, G.C., and A.D. Karayiannis. 2010. Morality, institutions and the wealth of nations: Some lessons from ancient Greece. *European Journal of Political Economy* 26: 68–81.

Black, D. 1948. On the rationale of group decision making. *Journal of Political Economy* 56: 23–34.

Black, D. 1963. *The theory of committees and elections*. Cambridge: Cambridge University Press.

Boland, P.J. 1989. Majority systems and the Condorcet jury theorem. *The Statistician* 38: 181–189.

Condorcet. 1785. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*, Paris: L'Imprimerie Royale.

Condorcet. 1989. *The political theory of Condorcet*. Trans. Fiona Sommerlad and Ian McLean (manuscript).

Faccarello, Gilbert. 2016. Marie-Jean-Antone-Nicolas Cariat de Condorcet (1743–1794). In *Handbook on the history of economic analysis*, vol. 1, ed. G. Faccarello, and H. Kurz. Cheltenham: Edward Elgar, forthcoming.

Grofman, B., and S.L. Feld. 1988. Rousseau's general will: A Condorcetion perspective. *American Political Science Review* 82: 567–576.

Grofman, B., G. Owen, and S.L. Feld. 1983. Thirteen theorems in search of the truth. *Theory and Decision* 15: 261–278.

Holler, M.J. 1980. What is paradoxical about the voting paradox. *Quality & Quantity* 14: 679–685.

Holler, M.J. 1982. The relevance of the voting paradox: A restatement. *Quality & Quantity* 16: 43–53.

Holler, M.J. 1994. Regulatory policymaking in a parliamentary setting: Comment. In *Jahrbuch für Neue Politische Ökonomie*, vol. 13, ed. Ph. Herder-Dorneich, K.-H. Schenk, and D. Schmidtchen. Tübingen: Mohr-Siebeck.

Holler, M.J. 2010. The two-dimensional model of jury decision making. *Journal of Public Finance and Public Choice (Economia delle Scelte Pubbliche)* 28(1): 29–42, published in 2012.

Holler, M.J., and S. Napel. 2007. Democratic decision procedures, stability of outcome, and agent power, with special reference to the European Union. In *Public choice and the challenges of democracy*, ed. J.C. Pardo, and P. Schwartz, 220–234. Cheltenham: Edward Elgar.

Kaniovski, S. 2008. Straffin meets Condorcet: What can a voting power theorist learn from a jury theorist? *Homo Oeconomicus* 25(2): 181–202.

Kaniovski, S. 2010. Aggregation of correlated votes and Condorcet's Jury Theorem. *Theory and Decision* 69(3): 453–468.

Kaniovski, S., and Zaigraev, A. 2011. Optimal jury design for homogeneous juries with correlated votes. *Theory and Decision* 71(4): 439–459.

Ladha, K.K. 1993. Majority-rule voting with correlated votes. *Social Choice and Welfare* 24: 69–85.

Laver, M., and K. Shepsle. 1990. Coalitions and cabinet government. *American Political Science Review* 84: 873–890.

McNutt, P. 2002. *The economics of public choice*, 2nd ed. Cheltenham: Edward Elgar.

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

Rothschild, E. 2001. *Economic sentiments: Adam Smith, Condorcet and the enlightenment*. Cambridge and London: Harvard University Press.

Saari, D.G. 1995. *Basic geometry of voting*. Berlin Heidelberg and New York: Springer.

Saari, D.G. 2011. 'Ranking wheels' and decision cycles. *Homo Oeconomicus* 28(3): 233–263.

Steunenberg, B. 1994. Regulatory policymaking in a parliamentary setting. In *Jahrbuch für Neue Politische Ökonomie*, vol. 13, ed. P. Herder-Dorneich, K.-H. Schenk, and D. Schmidtchen. Tübingen: Mohr-Siebeck.

Wittman, Donald A. 1973. Parties as utility maximizers. *American Political Science Review* 67: 490–498.

Updating the Law and Economics of Legal Parochialism

Nuno Garoupa

1 Introduction

I have known Jürgen Backhaus since my early days in my career. One of the many issues he has discussed and analyzed in his scholarship is the possible German origins of Law and Economics, that is, the use of economic arguments and reasoning when understanding the law in the German tradition. Modern Law and Economics is based on the neoclassical paradigm, so a recent development from the viewpoint of legal history. However, the pre-neoclassical Law and Economics of imminent European origin has had limited influence in current Law and Economics dominated by American oriented scholarship. It seems to me that such observation underlies a broader problem, which I have defined as "legal parochialism."

The difficult path of Law and Economics outside of the United States has been a matter of debate for many years. When the article on why Law and Economics seems to fail outside of the United States was published[1] in 2008, we did not predict the extent to which the discussion would raise so much interest and attract so many

This is a deeply revised version of an original article "The Law and Economics of Parochialism," University of Illinois Law Review (2011). The usual disclaimers apply.

[1]Nuno Garoupa & Thomas S. Ulen, *The Market for Legal Innovation: Law and Economics in Europe and the United States*, 59 ALA. L. REV. 1555 (2008).

N. Garoupa (✉)
Texas A&M University, Fort Worth, TX, USA
e-mail: nunogaroupa@law.tamu.edu

A. Marciano and G.B. Ramello (eds.), *Law and Economics in Europe and the U.S.*,
The European Heritage in Economics and the Social Sciences,
DOI 10.1007/978-3-319-47471-7_10

authors to propose alternative explanations.[2] Although there is disagreement over what actually explains the skeptical reception of Law and Economics outside of the United States, the entire body of literature provides systematic evidence of the following well-known facts: Law and Economics is influential in American and Israeli legal scholarship but it has little impact elsewhere; Law and Economics is dominated by legal scholars in the United States and in Israel but by economists elsewhere; the rate of acceptance of Law and Economics in American and Israeli courts is not impressive, but nevertheless significant, whereas the field is virtually ignored by courts elsewhere, albeit with some occasional references.

Obviously, there might be different degrees or tones of pessimism concerning these facts. Even in the United States, the alleged success and influence of Law and Economics is subject to different perspectives. For example, the established fact that some judges use Law and Economics in the federal courts could be seen as evidence of the importance of the field.[3] On the other hand, the fact that only some judges employ it could raise pessimistic concerns that the importance of Law and Economics in case law is still limited.[4] Several legal scholars have noted that Law and Economics is one of fastest growing fields in the United States.[5] At the same time, the numbers suggest that the field of Law and Economics is still composed of a minority of legal scholars largely confined to the top law schools.[6] In conclusion, in the United States, it is a matter of the half-empty and half-full bottle. But this is definitely not the case outside of the United States and Israel. There, we have a case of the fully empty bottle.

---

[2]See, e.g., Kenneth G. Dau-Schmidt & Carmen L. Brun, *Lost in Translation: The Economic Analysis of Law in the United States and Europe*, 44 COLUM. J. TRANSNAT'L L. 602 (2006); Alan J. Devlin, *Law and Economics*, 46 IRISH JURIST 2 (2011); Oren Gazal-Ayal, *Economic Analysis of "Law and Economics"*, 35 CAP. U. L. REV. 787 (2007); Oren Gazal-Ayal, *Economic Analysis of Law in North America, Europe and Israel*, 3 REV. L. & ECON. 485 (2007); Kristoffel Grechenig and Martin Gelter, *The Transatlantic Divergence in Legal Thought: American Law and Economics vs. German Doctrinalism*, 31 HASTINGS INT'L and COMP. L. REV. 295 (2008); Kilian Reber, *Once Upon a Time in America: Barriers to the Diffusion of Law and Economics* (July 17, 2008) (unpublished manuscript), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1161129.

[3]See, e.g., Stephen J. Choi and G. Mitu Gulati, *Mr. Justice Posner? Unpacking the Statistics*, 61 N. Y.U. ANN. SURV. AM. L. 19, 20–21 (2005).

[4]Garoupa and Ulen, *supra* note 1, at 1574.

[5]See, e.g., NEIL DUXBURY, PATTERNS OF AMERICAN JURISPRUDENCE 416–17 (1997); Robert Ellickson, *Trends in Legal Scholarship: A Statistical Study*, 29 J. LEGAL STUD. 517, 524–25 (2000); William M. Landes & Richard A. Posner, *The Influence of Economics on Law: A Quantitative Study*, 36 J.L. & ECON. 385, 385 (1993); Richard A. Posner, *Legal Scholarship Today*, 115 HARV. L. REV. 1314, 1316–17 (2002); Richard A. Posner, *The Sociology of the Sociology of Law: A View from Economics*, 2 EUROPEAN J. L. & ECON. 265, 275 (1995); Thomas S. Ulen, *A Crowded House: Socioeconomics (and Other) Additions to the Law School and Law and Economics Curricula*, 41 SAN DIEGO L. REV. 35, 35–37 (2004); Thomas S. Ulen, *A Nobel Prize in Legal Science: Theory, Empirical Work, and the Scientific Method in the Study of Law*, 2002 U. III. L. REV. 875, 906–07; Thomas S. Ulen, *The Impending Train Wreck in Current Legal Education: How We Might Teach Law as the Scientific Study of Social Governance*, 6 UNIVERSITY OF ST THOMAS LAW JOURNAL 303 (2009).

[6]Garoupa & Ulen, *supra* note 1, at 1573–74.

For many years, the argument was that Europe, Asia, and Latin America were moving toward neoclassical Law and Economics, but at a slower pace than the United States. That may well be the case, since it is likely that a window of time of 40 or 50 years is not enough to convincingly reject such a hypothesis. At the very least, however, we can say it is taking a long time. There is more neoclassical Law and Economics now in Europe, Asia, and Latin America than ever.[7] The influence of this work in legal scholarship and in legal policy is overwhelmingly disappointing, however, probably with the exceptions of competition and corporate law. As we noted in our original article, in vivid contrast with the United States, the main textbooks and treatises on property, contract, and tort law, or on procedure outside of the United States, simply ignore Law and Economics insights.[8]

Naturally, the lack of success of Law and Economics has intrigued scholars. The first wave of explanations developed in the early 1990s, still hoping that the whole process was merely delayed, focused on the obvious differences between the United States and the rest of the world. The most obvious and most popular explanations, for a while, were legal tradition (civil law is over-theorized, whereas common law is under-theorized) and language (legal scholars do not read English).[9] The problem is that these popular explanations seemed to neglect that Law and Economics also was not popular in the United Kingdom, nor in any other common law jurisdiction (with the exception of Israel, as we already noted).[10] In fact, looking at current trends in Europe, Law and Economics is in much better shape in Italy, Germany, and Spain than in the United Kingdom or in Ireland.[11] The idea that Law and Economics is

[7]There have been annual meetings of the European Association of Law and Economics since the early 1980s, the Latin American Law and Economics Association (ALACDE) since the mid-1990s, and the Asian Law and Economics Association since 2005.

[8]Garoupa & Ulen, *supra* note 1, at 1575–76.

[9]See, e.g., Brian R. Cheffins, *The Trajectory of (Corporate Law) Scholarship*, 63 CAMBRIDGE L.J. 456, 461–62 (2004); R. Cooter & J. Gordley, *Economic Analysis in Civil Law Countries: Past, Present, and Future*, 11 INT'L REV. L. & ECON. 261, 262 (1991); Dau-Schmidt & Brun, *supra* note 2, at 617–18; Aristides N. Hatzis, *The Anti-Theoretical Nature of Civil Law Contract Scholarship and the Need for an Economic Theory*, 2 COMMENTS. ON L. & ECON. 1, 30–31 (2002); Richard A. Posner, *Law and Economics in Common-Law, Civil-Law, and Developing Nations*, 17 RATIO JURIS 66, 76–77 (2004); Richard A. Posner, *The Future of the Law and Economics Movement in Europe*, 17 INT'L REV. L. & ECON. 3, 4–5 (1997); Hans-Bernd Schäfer, *What Are the Practical Implications of Law and Economics Research in Germany?*, *in* PETER NOBEL AND MARINA GETS EDS., NEW FRONTIERS IN LAW AND ECONOMICS (2006).

[10]See, e.g., Christopher McCrudden, *Legal Research and the Social Sciences*, 122 L.Q. REV. 632, 639 (2006); A.I. Ogus, *Law and Economics in the United Kingdom: Past, Present, and Future*, 22 J.L. SOC'Y 26, 29–30 (1995).

[11]Recent developments include the annual meetings of the German Law and Economics Association (GLEA) since 2003, the Italian Society of Law and Economics (SIDE) since 2005, the Spanish Law and Economics Association (AEDE) since July 2010, and more recently the Polish Association of Law and Economics (PSEAP) since October 2012. Outside of Europe, we should notice the Brazilian Association of Law and Economics (ABDE) with meetings since October 2008. No such associations exist in the United Kingdom or in Ireland.

more relevant for judge-made law than for a legal system dominated by codification is simply rejected by the evidence.

A second popularized account was the need for a Legal Realism revolution to precede the expansion of Law and Economics.[12] Such explanation only pushes the discussion backwards, however, to a previous step: why was Legal Realism successful in U.S. legal academia but not elsewhere? In fact, it is accepted that Legal Realism did not even originate in the United States.[13]

Another explanation looked at ideology and legal philosophy.[14] The mistake here was to think of Law and Economics as a mere conservative legal movement to reduce the influence of liberals in law schools.[15] In our original article, we do not deny that Law and Economics was financed by conservative foundations, nor that it attracted many conservative legal scholars.[16] Nevertheless, there are many liberals producing excellent Law and Economics scholarship. Furthermore, while Law and Economics was initially associated with the University of Chicago (still a major player in the field nowadays), it has now emerged in many law schools that cannot be labeled as conservative. At the same time, if it is all about a particular ideology or a certain judicial philosophy, we would expect the liberal (or perceived to be liberal) legal movements to be very successful in Europe and elsewhere. As I argued I my original Article, this is hardly the case. In fact, my own perception is that, in Europe and in Asia, unlike in the United States, law schools are not seen as more liberal than economic departments or business schools. Yet, according to this explanation, pro-market theories popular with European, Latin American, and Asian economists were rejected by European, Latin American, and Asian legal scholars. It does not seem to be a convincing explanation.

We should acknowledge that legal scholars in Europe show an intense dislike for efficiency and seem to be much more open to social justice or redistributive legal arguments.[17] Chronologically, however, the distaste for efficiency seems to have been revealed when confronted with neoclassical Law and Economics. Therefore, it is unclear whether a neoclassical approach to law is rejected because legal scholars

[12]See, e.g., Charles K. Rowley, *An Intellectual History of Law and Economics: 1739–2003*, *in* THE ORIGINS OF LAW & ECON.: ESSAYS BY THE FOUNDING FATHERS 3, 11–12 (Francesco Parisi & Charles K. Rowley, eds., 2005).

[13]*But* See Steven G. Medema, *Wandering the Road from Pluralism to Posner: The Transformation of Law and Economics*, 3 HIST. POL. ECON. (SUPPLEMENT) 202, 204 (1998); Rowley, *supra* note 12, at 3–29.

[14]Dan-Schmidt & Brun, *supra* note 2, at 616; Garoupa & Ulen, *supra* note 1, at 1578–79.

[15]See the debate by several scholars in *Symposium: Calabresi's* The Costs of Accidents*: A Generation of Impact on Law and Scholarship*, 64 MD. L. REV. 1 (2005), in particular, articles by Adam Benforado & Jon Hanson, *The Costs of Dispositionism: The Premature Demise of Situationist Law and Economics*, 64 MD. L. REV. 31 (2005); Anita Bernstein, *Whatever Happened to Law and Economics?*, 64 MD. L. REV. 303 (2005); and Ugo Mattei, *The Rise and Fall of Law and Economics: An Essay for Judge Guido Calabresi*, 64 MD. L. REV. 220 (2005).

[16]See Garoupa & Ulen, *supra* note 1, at 1579–82.

[17]See Catherine Valcke, *The French Response to the World Bank's* Doing Business *Reports*, 60 U. TORONTO L.J. 197 (2010).

dislike efficiency, or efficiency is disliked because legal scholars rejected Law and Economics. It is clear that the discussion about efficiency as a relevant normative criterion for law has not been fruitful in Europe, where the overwhelming majority of legal scholars have joined the "no" side. This discussion did not take place, however, until European legal scholarship was confronted with the developments of neoclassical Law and Economics. So the explanation that Law and Economics has been rejected by traditional legal scholarship in Europe because efficiency (or any utilitarian criteria) is philosophically inconsistent with European legal thought seems too simplistic, and, to a certain extent, naïve. On a different matter, this explanation has also been used by Anglo American audiences to justify why common law is efficient and civil law is inefficient, as if efficiency of the law varies with the philosophical beliefs held by law professors.[18]

The failure of these early explanations called attention to a basic insight: the world of legal thinking is more complex! Legal culture, language, ideology, and legal philosophy are important factors that explain the particularities of legal scholarship in different countries. Altogether, however, they fail to pin down why Law and Economics has been so unsuccessful outside of the United States. Moreover, in a globalized world that has reduced cultural barriers, and when English has assumed the role of *lingua franca*, the prediction in the early 1990s was that soon Law and Economics would enjoy the same popularity in and outside of the United States.[19] Obviously, these predictions turned out to be largely incorrect.

By the mid-2000s, legal scholars produced a second-wave of explanations. These explanations reflected the struggle Law and Economics experienced outside of the United States while booming there and in Israel. They included the weak economic training of lawyers in Europe, in Latin America, and in Asia (no mathematical background[20]), the incentives established by legal academia in Europe that largely favor conformity rather than innovation (the core argument of my original article with Thomas Ulen[21]), or the relevance of a start-up process of legal innovations (concerning the relevance of publication outlets[22]). In my view, the important contribution of this second-wave literature is to expose Law and Economics as a legal innovation, and legal innovations can only be produced in competitive markets that generate appropriate incentives. If the incentives are appropriately designed, Law and Economics flourishes; if not, Law and Economics has a hard time emerging as a meaningful player in legal scholarship.[23]

---

[18]See Nuno Garoupa & Carlos Gómez Ligüerre, *The Syndrome of the Efficiency of the Common Law*, 29 B.U. INT'L L. J. 287 (2011).

[19]See Cooter & Gordley, *supra* note 9, at 261–63.

[20]*Contra* Anthony Ogus, *Law and Economics in the Legal Academy, or, What I Should Have Said to Discipulus*, 60 U. TORONTO L.J. 169, 170 (2010); Dennis W.K. Khong, *On Training Law and Economics Scholarship in the Legal Academia*, 1 ASIAN J. L. & ECON., No. 2, 2010.

[21]Garoupa & Ulen, *supra* note 1, at 1603–04.

[22]Reber, *supra* note 2, at 12–15.

[23]Gazal-Ayal, Economic Analysis of "Law & Economics," *supra* note 2, at 787–98.

In the original article, we discuss in detail the development of the adequate incentives for legal innovation to flourish.[24] We mention the rigidity of the hierarchical relationships established in legal academia as a major drawback (whereas in the United States we tend to have an inverted pyramid—that is, many full professors and few untenured faculty—in Europe, we have a regular pyramid—that is, many untenured professors and few senior faculty).[25] In Europe, there is no meaningful academic mobility (in particular, no lateral mobility since salaries are rigid and usually fixed by the government). Law reviews are faculty-edited and value methodological conformity. The market for new law professors promotes inbreeding and the entrenchment of the *status quo*. We can easily extend our description to Asia and Latin America.

All of these explanations are important and relevant, but they do not seem to capture the full puzzle. Somehow they seem to be circular. If incentives explain the success and failure of Law and Economics across jurisdictions, a new question emerges: why is it that some jurisdictions seem to be able to develop the adequate incentives to promote legal innovation, whereas others support and protect the *status quo*? There is always the possible explanation of chance: the United States and Israel were lucky, whereas the rest of the world was unlucky. In the original Article, I do not exclude the possibility of the great man or woman theory, wherein a famous champion of a particular legal innovation promotes its development through legal entrepreneurship.[26] Such an explanation, however, does not seem to be accurate as a general basis for legal innovations.

In fact my aim is to emphasize a more general perspective. In particular, I suggest that there is nothing particular to Law and Economics that explains its current lack of success outside of the United States. In my view, this is related to a more general problem, which I call "legal parochialism." My argument was already echoed in my earlier article with Thomas Ulen, but it was largely neglected in the debate.[27] I think, however, that this legal parochialism is a more important insight proposed in the original article than the competitiveness of the market for legal scholarship, the latter being much more noticed by other authors.

This Article proceeds in the following order. Part II explains the general problem in more detail. Part III discusses legal parochialism. Part IV concludes the paper.

[24]Garoupa & Ulen, *supra* note 1, at 1627–31.

[25]*Id.* at 1603–04.

[26]*Id.* at 1611–14.

[27]*Id.* at 1632.

2 A More General View

The explanations for why Law and Economics has largely failed outside of the United States neglect the fact that other areas of the "Law and" movement have been similarly unsuccessful abroad, most of them with little or no influence from economics.[28] Consider all the following fields of legal scholarship: empirical legal studies,[29] critical legal studies,[30] feminist jurisprudence,[31] law and literature,[32] law and politics,[33] law and psychology,[34] law and cognitive sciences,[35] law and biology,[36] and law and anthropology.[37] All of them are important American legal innovations of the last decades. They all have been largely ignored outside of the United States. Clearly, lack of success outside of the U.S. legal academia has little to do with Law and Economics.

It could be that all these legal innovations were received with skepticism outside of the United States because these were developments taking place in the United States. My argument, however, is that such skepticism is not particular to American legal products. While Law and Economics became popular in the United States, other legal innovations were developed around the world. Here are a couple of

---

[28]Marc Galanter & Mark Alan Edwards, *Introduction: The Path of the Law* Ands, 1997 WIS. L. REV. 375, 381.

[29]For examples, see generally issues of the *Journal of Empirical Legal Studies*, as well as Theodore Eisenberg, *The Origins, Nature, and Promise of Empirical Legal Studies and a Response to Concerns*, 2011 U. LLL. L. REV. 1713 (2011).

[30]See, e.g., MARK KELMAN, A GUIDE TO CRITICAL LEGAL STUDIES (1987); ROBERTO MANGABEIRA UNGER, THE CRITICAL LEGAL STUDIES MOVEMENT (1986).

[31]See, e.g., JUDITH A. BAER, OUR LIVES BEFORE THE LAW: CONSTRUCTING A FEMINIST JURISPRUDENCE (1999); FEMINIST JURISPRUDENCE (Patricia Smith ed., 1993); Gillian K. Hadfield, *Feminism, Fairness, and Welfare: An Invitation to Feminist Law and Economics*, 1 ANN. REV. L. & SOC. SCI. 285 (2005); Ann C. Scales, *The Emergence of Feminist Jurisprudence: An Essay*, 95 YALE L.J. 1373 (1986).

[32]See, e.g., RICHARD A. POSNER, LAW AND LITERATURE (1998); LAW AND LITERATURE: CURRENT LEGAL ISSUES 2 (Michael Freeman & Andrew D.E. Lewis eds., 1999).

[33]See generally issues of the *Journal of Law and Politics* (University of Virginia) and the *Texas Review of Law and Politics* and more recently *Journal of Law and Courts*, as well as Keith E. Whittington et al., *The Study of Law and Politics*, *in* THE OXFORD HANDBOOK OF LAW AND POLITICS (Keith E. Whittington et al., eds., 2008).

[34]See generally issues of *Law and Psychology Review* and *Psychology, Public Policy and the Law* as well as BEHAVIORAL LAW AND ECONOMICS (Cass R. Sunstein, ed., 2000) and Christine Jolls, *Behavioral Economics and Its Applications*, *in* BEHAVIORAL LAW AND ECONOMICS 115 (Peter Diamond & Hannu Vartiainen eds., 2007) (providing an overview of the field).

[35]See, e.g., Matthias Mahlmann & John Mikhail, (2005).

[36]See, e.g., Owen D. Jones & Timothy H. Goldsmith (2005); Owen D. Jones, (2001).

[37]ANTHONY G. AMSTERDAM & JEROME BRUNER, MINDING THE LAW (2000); JOHN M. CONLEY & WILLIAM M. O'BARR, JUST WORDS: LAW, LANGUAGE, AND POWER (2005); BETWEEN LAW AND CULTURE: RELOCATING LEGAL STUDIES (David Theo Goldberg et al., eds, 2001); .INSIDE AND OUTSIDE THE LAW: ANTHROPOLOGICAL STUDIES OF AUTHORITY AND AMBIGUITY (Olivia Harris, ed. 1996.)

examples: Socio-legal studies in the United Kingdom[38]; *Droit Économique* (Economic Law) and *Sociologie du Droit* (Sociology of Law) in France[39]; and *Wertungsjurisprudenz* (Jurisprudence of Value Judgments, used in interpretation of private law), *Verwaltungswissenschaft* (Science of Administration) and *Staatswissenschaften* (German Law and Politics) in Germany.[40] All these innovations in legal thinking had little to no influence in the United States. In fact, most of these innovations had little impact outside of the legal culture in which they were developed.

Some of these foreign innovations closely trace American innovations. For example, U.S. law and society, U.K. socio-legal studies, and French *Sociologie du Droit* can be regarded as adjacent developments in legal scholarship. There are some important differences, but we could argue that they do not constitute major impediments to scientific dialog. The British version was initially more empirically oriented, whereas the French version has been more theorized, and the American version is probably somewhere in between. Nevertheless, these legal methodologies largely share the same concerns and legal perspectives. Still, a quick look at the main journals of these fields tells us that there are few cross-references.[41]

Other innovations reveal the polarization of legal debate, whereas Law and Economics focuses on efficient legal rules based on the aggregation of individual preferences and actions, *Droit Économique* studies the mechanisms adequate to favor state intervention from the perspective of the interests of the state. Law and Economics frames state intervention in the context of market failures. *Droit Économique* rejects the role of the market altogether. Not surprisingly, broadly speaking, each side ignores the others.

Furthermore, with the exception of German legal science from the early 1900s, foreign developments of legal scholarship have been virtually ignored in the United States, even those in English-speaking countries with a common law tradition.[42]

---

[38]See generally the *Socio-Legal Studies Association* and the issues of the *Journal of Law and Society* and *Social & Legal Studies*, as well as A. JAVIER TREVINO, THE SOCIOLOGY OF LAW (2007) (summarizing the vast literature on socio-legal studies produced in the United Kingdom and other commonwealth jurisdictions).

[39]For examples, see generally the *Association Internationale de Droit Économique* and the issues of the *Revue Internationale de Droit Économique* as well as LAURENCE BOY ET AL., DROIT ÉCONOMIQUE ET DROITS DE L'HOMME (2009) and JEAN-PAUL VALETTE, DROIT PUBLIC ECONOMIQUE (2009) (providing general introductions to the field). Both French methods were deeply influenced by Marxism and Marxist legal scholars.

[40]*See also* REINHOLD ZIPPELIUS, INTRODUCTION TO GERMAN LEGAL METHODS (2008) (explaining German law and summarizing methodological developments in German legal scholarship).

[41]Many legal scholars of these movements attend the annual meeting of the Law and Society Association yet they seem to develop their legal methods largely independently. See generally issues of the *Journal of Law and Society* and *Law and Society Review*, as well as KITTY CALAVITA, INVITATION TO LAW AND SOCIETY: AN INTRODUCTION TO THE STUDY OF REAL LAW (2010) (summarizing the American-based law and society scholarship) and MARK KELMAN, *supra* note 30.

[42]David S. Clark, *Development of Comparative Law in the United States*, *in* THE OXFORD HANDBOOK OF COMPARATIVE LAW 187 (Mathias Reimann and Reinhard Zimmermann, eds., 2006).

The explanation I propose for these facts is legal parochialism. Law and Economics, as well as other developments of the "Law and" movement, has been a victim of legal parochialism outside of the United States inasmuch as other legal innovations have been victims of legal parochialism inside the United States. In my view, legal parochialism provides a more comprehensive understanding of the problem. Most of the first- and second-wave explanations simply reflect different aspects of this significant legal parochialism.

3 Legal Parochialism

My theory is that legal parochialism operates like protectionism in trade.[43] We can envisage a global market for legal innovations in scholarship. Each jurisdiction protects its own market from foreign competition. The protection of the local market is important for local producers (i.e., legal scholars). It increases their return on local human capital, including providing for important network effects.[44] Therefore, they have an interest in avoiding foreign competition.

The first condition to implement protectionism in legal innovations is to recognize that a small group of producers is able to cartelize and secure the benefits from closing down the market. At the same time, the losers (the local consumers) must have dispersed interests that can hardly be coordinated to avoid protectionism.[45] So legal parochialism emerges as the result of cartel behavior from the main beneficiaries: the local incumbents, who are the law professors and legal scholars.

The second condition is that the local incumbents exercise some significant market power in the relevant industry. Such market power can be reflected in different relevant dimensions. For example, they can exert some control over the supply of legal reform so that legal policymakers do not look for alternative providers elsewhere. In this respect, local incumbents operate as a powerful lobby that discourages legal reformers from looking for multiple sources of legal thinking.

At the same time, local incumbents will push for "subsidizing" local production using arguments that are similar to the traditional infant industry rationale.[46] Typically, they will insist on culture, language, history, and other national symbols to promote protection and allege a better understanding of the local needs. Within this protectionist view, potential competition has to be excluded in order for the

[43]For an introduction, see KYLE BAGWELL & ROBERT W. STAIGER, THE ECONOMICS OF THE WORLD TRADING SYSTEM (2002), and ROBERT C. FEENSTRA, ADVANCED INTERNATIONAL TRADE: THEORY AND EVIDENCE (2004).

[44]See Nuno Garoupa & Anthony Ogus, *A Strategic Interpretation of Legal Transplants*, 35 J. LEGAL STUD. 339 (2006); Anthony Ogus, *The Economic Basis of Legal Culture: Networks and Monopolization*, 22 OXFORD J. LEGAL STUD. 419 (2002).

[45]See FEENSTRA, *supra* note 43, at 300–337.

[46]*Id.*

local incumbents to keep additional rents. Rejection of foreign law, sources, legal education, and legal practice is promoted to avoid market contestability.

Legal parochialism, in my view, is just a form of trade protectionism in the context of the market for legal ideas. The consequences of legal parochialism are the standard losses from trade protectionism: less efficient allocation of resources in supply of legal norms, under-development of new legal innovations, and significant opportunity costs disseminated across society.[47]

Legal parochialism is stronger in some jurisdictions and weaker in others, depending on the combination of market determinants. A larger local incumbent profession reduces the cohesion and ability for cartel behavior. Consequently, not only is legal parochialism somehow diluted, but also the large size of the local legal profession creates the conditions for a competitive market, thus significantly reducing the costs of protectionism. On the contrary, a smaller local incumbent profession supports a more consistent cartel and decreases competition in the local market, therefore enhancing the costs of protectionism.

My perception is that the United States is an example of a large local incumbent profession, whereas most European, Latin American, and Asian jurisdictions are examples of the opposite case. I do not suggest that the United States is an exception; quite the contrary. I think the United States is consistent with the model, but because the market determinants are different, due to the size of the legal profession, legal parochialism is likely to be weaker. As recognized by the literature on trade protectionism, however, notice that legal parochialism is consistent with an aggressive strategy of exporting legal scholarship. Educating foreign human capital, establishing an international network of legal thinking, and influencing foreign legal reforms are not in contradiction with legal parochialism, whereas legal parochialism is about deterring importation of legal ideas, educating foreign human capital, and influencing foreign legal reform are about exporting legal ideas.

Another example I have mentioned is Israel. Again, I do not think this is an exception to my model. They have been in the early stages of the process of shaping their local legal system and so, presumably, they have been in a situation where the incumbent local cartel is relatively weak. As a consequence, legal parochialism has not been able to exert the same influence in Israel as it has elsewhere. As for the future, I predict that legal parochialism will be stronger in Israel in the decades to come as the process of shaping their legal system reaches maturity.[48] At the same time, I suggest that the fairly positive reception of Law and Economics, and other American legal innovations, in small legal communities like Taiwan reflects very much the same lack of a well developed and established incumbent legal theory.[49]

---

[47] *Id.*

[48] See Haim Sandberg, *Legal Colonialism—Americanization of Legal Education in Israel*, 10 GLOBAL JURIST, MAR. 2010.

[49] For a more general discussion of successful American legal education in Asia, see Gail J. Hupper, *The Academic Doctorate in Law: A Vehicle for Legal Transplants?*, 58 J. LEGAL. EDUC. 413 (2008).

Presumably, we should expect an identical course of action in mainland China in the next decade or so.

Keeping the metaphor between legal parochialism and trade protectionism, we can identify serious threats to legal parochialism in recent years. Presumably, these serious threats will increase in the coming decades. As with protectionism in trade, legal parochialism can be significantly reduced by pressure of external forces. The most obvious one is the globalization of legal services that has resulted in the ongoing globalization of legal education. In the United States, the importance of international, comparative, and foreign law in the J.D. curriculum is still disappointing, but most law professors would recognize that offerings of such courses have increased steadily in the last decades. My impression is that a similar path is recognizable in Europe. International exchange of students and faculty in law schools is still probably insignificant compared to the hard sciences and the social sciences. But the general impression is that exchange programs have more demand now than they did years ago.

A second threat to legal parochialism has been the integration of legal markets. Such process has had remarkable effects in Europe, as national law subsided to European Union law in many relevant fields. There have been occasional backlashes exhibiting strong legal parochialism, but the process has eroded significantly the power of local legal elites and has forced the open exchange of ideas in Europe.[50] Obviously, we are still far from a fully integrated market for legal ideas in Europe, as some barriers are artificially kept by local incumbents, but progress in the last decades is noticeable.

Another example of integration of legal markets is the explosion of legal transplants promoted by the governments of developed countries and international organizations. Such movement had two important consequences for the market of legal ideas. Many legal innovations are now competing for the market of legal ideas in developing economies, with mainland China being an obvious example. The United States and Europe are aggressively exporting their legal models, educating foreign human capital, and lobbying legal reform. The expansion of American and European law schools to Asia in the last decade is remarkable. Legal ideas and traditions are forced to compete in distant markets. At the same time, they are also pushed to compete inside international organizations in order to capture and determine their legal policy agendas. In this respect, for example, law and economics has been remarkably successful due to the strong position of economists in international organizations such as the World Bank or the International Monetary Fund.[51]

---

[50]For example, as mentioned by Garoupa and Ulen, consider the following episode: "[I]n a letter to the President of France, [in December of 2006,] forty well-known French law professors … rejected EU law as law that deserves to be studied and analyzed. Only French law should be taught at French law schools, according to these law professors." Garoupa & Ulen, *supra* note 1, at 1626 n.313.

[51]See Galit A. Sarfaty, *Why Culture Matters in International Institutions: The Marginality of Human Rights at the World Bank*, 103 AM. J. INT'L L. 647 (2009).

4 Conclusions

The main thesis of this Article is that the slow growth of law and economics outside of the United States is part of a more general problem: legal parochialism. At the same time, the effects of legal parochialism can be less severe if the "protected market" is significantly large, with the United States fitting this particular case. Conversely, if the "protected market" is small and easily cartelized, the effects of legal parochialism can be important, with Europe, Latin America, and Asia exemplifying this situation more consistently.

The challenging note is always to explain the difficult reception of Law and Economics outside of the United States, whereas neoclassical economics has become dominant around the world. More generally, in the context of my theory, there must be an explanation for why we have legal parochialism, and not parochialism in other social sciences (economics, sociology, and psychology), that is, in fields of study where the argument of local culture to justify isolationism could apply as well. In my view, the reason is purely market-driven and explained by non-academic rents. Scholars in economics, sociology, and psychology, and even business studies, do not significantly derive their income from sources outside of academia. Naturally, there is less concern in protecting the domestic market because the demand for rent-seeking is less important (albeit not absent). Law professors outside of the United States make most of their income outside of academia, strictly speaking, as they practice law, advise the government, and play an active role in lawmaking (for example, they tend to dominate code redrafting committees or law commissions).[52] These outside activities generate significant rents; thus, protecting the domestic market is of utmost importance.[53] Therefore, my explanation does not rely on law being different from other social sciences, but on market opportunities.

On a positive note, I share the optimism of many legal scholars who have predicted the expansion of Law and Economics outside of the United States. By recognizing legal parochialism, I suggest that it is easier to understand that the future of Law and Economics outside of the United States requires a careful focus on local legal problems. Such a path requires excellent comparative work, an application of Law and Economics to local legal doctrines, and local publication outlets.[54]

---

[52]These rents vary across fields of law. Presumably business law is more profitable than legal history.

[53]This effect could be reinforced if the outside market is dominated by legal scholars from the top universities who have no interest in sharing their rents with other legal scholars from lower ranked schools, let alone scholars with foreign methodologies or foreign legal education. In fact, if the outside market is dominated by a small handful of top law professors, politically and socially influent, the ideal conditions for cartel behavior are more likely to be satisfied.

[54]Such as the *European Journal of Law and Economics*, the *Asian Journal of Law and Economics* and the new *Latin American Journal of Law and Economics*.

More recently, we have followed the development of empirical legal studies.[55] Once more, as with Law and Economics, its expansion and acceptance outside of the United States has been slow and limited. The Society of Empirical Legal Studies is massively American and there are no equivalent organizations elsewhere. The Journal of Empirical Legal Studies is dominated by American scholars with occasional papers making use of non-American datasets. Yet empirical legal studies do not suffer from the standard critiques. They are unrelated to efficiency or to any alleged conservative movement. Datasets are now widely available and funding for assembling data is not difficult to find outside of the United States. It does seem to me that once again "legal parochialism" plays a significant role that delays the application of empirical legal studies to relevant legal problems outside of the United States.[56]

[55]Supra note 29.

[56]See Emanuel V. Towfigh, *Empirical Arguments in Public Law Doctrine: Should Empirical Legal Studies Make a Doctrinal Turn?*, 12 INT'L J. CON. L. 670 (2014).

Procedure

Peter Lewisch and Jeffrey Parker

1 Introduction

(a) In its broadest sense, procedural law is the "law of law". Procedural law is concerned with the making, application, and enforcement of substantive law. In this broad sense, procedural law would include constitutional rules of law production, such as constitutional arrangements among the branches of government or criteria for the enactment of legislation and administrative regulations. In a narrower sense, procedural law is concerned with specifying the conditions under which legal controversies are adjudicated by tribunals, including courts, administrative bodies, or other dispute-resolution entities such as arbitral panels. It is this narrower sense of procedural law to which this chapter is addressed primarily.

(b) Conventionally, the law governing the adjudication of legal disputes includes the description of what disputes are ripe for determination by adjudication, what tribunals have the competency or jurisdiction to consider the dispute, how the applicable rules of substantive law are ascertained and applied, how the facts pertaining to the dispute are investigated and found, how the decision or judgment is rendered, what effect that judgment may have on subsequent disputes, and how that judgment may be reviewed, attacked, or reexamined. Procedural rules also govern the enforcement of judgments.

(c) From the economic perspective, procedural law may be studied from either a macroscopic (aggregated) or a microscopic (individual cases) viewpoint. There is some literature studying procedure from the macroscopic viewpoint.

P. Lewisch (✉)
University of Vienna, Vienna, Austria
e-mail: Peter.lewisch@univie.ac.at

J. Parker
George Mason University, Fairfax, VA, USA
e-mail: Jparke3@gmu.edu

A. Marciano and G.B. Ramello (eds.), *Law and Economics in Europe and the U.S.*,
The European Heritage in Economics and the Social Sciences,
DOI 10.1007/978-3-319-47471-7_11

A recent empirical case study on Austria has found a significant positive correlation between the increase in the real GNP/person and the volume of litigation.[1] Most of the existing law-and-economics literature takes a microscopic viewpoint, focusing on the combined effects of procedural and substantive rules on the incentives and behaviors of individual participants in both civil and criminal cases, including the parties to an adjudication and the individuals composing the tribunal.

(d) From this viewpoint, the dominant economic model of legal procedure is what might be termed the "expected value" model of adjudication. This model explains litigation as the result of the predictions of both the plaintiff and defendant—actual or potential—as to the envisaged outcome in terms of the respective costs and benefits involved. Because in most such cases the existence of the legal entitlement is in dispute, the parties' expectations are deflated by the probability of success in the claim, and hence the term "expected value" when referring to contested claims. The "expected value" model poses the problem as one of reconciling the differing views of the parties to a legal dispute, and of characterizing the interactive effects of the parties' actions on each other. Therefore, a great deal of attention has been devoted to considering the conditions under which the parties' expectations may converge, thereby allowing the dispute to be settled, that is, to be resolved by formal agreement before definitive adjudication. However, this approach obviously presupposes some type of default outcome in the absence of agreement. Furthermore the adjudicated outcome is not simply a matter of distribution between contesting parties, but can have external effects on social welfare. The conventional point of view thus poses the social problem of adjudication as minimizing the sum of two types of costs: (i) the "direct" costs of the adjudication itself; and (ii) the "error" costs associated with legally or factually erroneous judgments, which can extend beyond the immediate parties. Even at this primitive level, it is clear that there is a trade-off between direct costs and error costs. However, one cannot fully characterize the costs of error without considering the underlying substantive legal rules that are sought to be enforced.

Thus, the economic analysis of procedural law does not draw such a sharp distinction between substance and procedure as do legal dogmatics. In the economic analysis of law, it is the combination of both substantive and procedural rules that determine the ultimate efficiency properties of the legal system. To illustrate, let us take the example from the law-and-economics literature of the distinction between a "property"-type rule and a "liability"-type rule. In essence, this distinction is one of remedy: a "property"-type rule is one

[1]See Clemenz and Gugler (2000). A second finding of this research was that the GNP growth rate per person is inversely related to the number of new law suits filed (per person), indicating that an economic boom correlates with a decrease of new litigation, while recession periods stimulate litigation. Note, however, that even on the basis of this empirical research the aggregated volume of litigation is ambiguous as to social welfare or efficiency. Despotic regimes may have little or no litigation, but this is not an indication of high social welfare.

enforced by specific order, while a "liability"-type rule is one enforced by money damages as ascertained by a tribunal. The relative efficiency properties of these two types of rules may depend upon procedural characteristics, such as whether economically compensable money damages may be determined by a tribunal within a tolerable range of direct and error costs. Note that such a decision may depend in part upon the properties of the procedural system as encouraging or discouraging strategic behaviors by the litigants, either in the course of the legal dispute or in the period before the legal dispute arose. If the procedural system is such as to discourage one or both parties from generating the necessary information ex ante, or revealing the specified datum ex post, then the characteristics of the procedural system may narrow the range of efficient substantive rules.

For this reason, it is not necessarily true that a publicly provided procedural system can be directed toward minimizing the costs of legal disputes, either to the parties or to the society as a whole. Reducing the costs of legal disputes too low will encourage an oversupply of disputes, while raising the costs of legal disputes too high will discourage legal interactions. Social welfare can be reduced by making legal disputes either too easy or too difficult. These problems are magnified when considering forms of legal dispute-resolution that deviate from the inter pares model of private civil litigation, such as criminal prosecution or other forms of public law enforcement.

(e) A more general consequence of this point is that procedural rules may be both complements of substantive law, as in conventional legal dogmatics, and substitutes for substantive rules. Thus, in a contractual situation, parties may be unable to agree on certain substantive terms (such as the price of an apartment to be purchased, or damages in the case of a contract breach), but may be able to agree in advance on a procedure that both would accept. This is the type of arrangement that is characteristic of private dispute-resolution arrangements, such as arbitration or, even more broadly, of negotiation-techniques. Whether parties will agree in advance to such arrangements depends in part on the properties of the procedural system that otherwise would apply.

(f) The remainder of this paper presents, first a general analysis of the incentives for privately and publicly provided law enforcement (Part 2) and then an economic analysis of civil and criminal procedure along the commonly employed legal desiderata of "just, speedy, and inexpensive" adjudication (Part 3).

2 Law Enforcement and the "Expected Value" Model

In this section, we present a general economic view of law enforcement, beginning with case of private enforcement and passing to public enforcement, including criminal law enforcement.

The micro-economic approach is individualistic. All social interaction is factored down to the choices of individual actors. Then, the incentives operating on those individual actors are examined to predict their actions. In its most general form, this analysis focuses on the "opportunity costs" of the individual. This means that the individual's incentives are compared with the next best use of the individual's time and effort. To the extent that the consequences of an individual's actions are felt by that individual, then the resulting costs are said to be "internalized" to that individual. To the extent that such consequences are not felt by that individual, then the resulting costs are said to be "externalized." As the reader is aware from other chapters, much of the law-and-economics literature is concerned with the significance of this distinction between "internalized" and "externalized" effects. The Coase theorem challenged the significance of the distinction per se, instead arguing that the efficiency of legal entitlements was more profoundly affected by the incidence of transaction costs on individuals' ability to reach socially optimal arrangements by private bargaining. Whereas the Coasian analysis suppressed the problem of costly enforcement of legal rights, this enforcement can be seen simply as another form of "transaction cost" standing in the way of optimal arrangements. Therefore, we can apply this same perspective to the analysis of law enforcement.

2.1 Private Enforcement

(a) From the economic point of view as interpreted by Coase, substantive rules of law assign private rights and obligations in order to induce optimal arrangements of resource use. Thus, as applied to ordinary civil litigation between private parties, substantive law assigns the starting allocations of private property rights. However, given that transaction costs are strictly positive in any actual case, the initial assignment of legal rights may influence social welfare. Furthermore, both enforcement costs and the form of remedy may influence the efficiency properties of the substantive law. If enforcement is too costly, then the initial assignment of property rights will have no effect, and would be equivalent to the non-assignment of rights, which is likely to be inefficient.

(b) Private enforcement of law seeks to align the enforcement incentives with the underlying substantive rule. Thus, the typical case is that the owner of the legal entitlement is given a right of action to enforce the entitlement against invasion by another (for example, the owner of a lot sues his neighbor to stop the strolling of his dog). To the extent that the owner-claimant "internalizes" the enforcement incentive by obtaining a private remedy, enforcement is obtained

in proportion to its social value, as assessed by the claimant (in the case of property-type rules) or as assessed by the court (in the case of liability-type rules). In these cases, the individual will provide law enforcement, if this enforcement is cost justified on the basis of her own costs and benefits, which are equivalent to social costs and benefits (= full convergence of private and social incentives to bring suit).

(c) However, there are several qualifications to this simplistic analysis, most tending toward the argument that private enforcement of law produces under-enforcement.

First, private and social incentives for enforcement may diverge. If individually provided law enforcement generates positive external effects to certain other individuals, to groups of individuals, or to the general public, these social benefits will not enter into the individual's cost calculus. Such an "external" benefit to society may stem from enforcement that either clarifies a legal rule or entitlement or deters third parties from committing similar invasions. In these cases, the individual is likely to systematically disregard those social benefits (that do not occur to herself) and to underprovide law enforcement. This argument often is said to justify a degree of public subsidy to private enforcement by using general tax revenues to cover most of the costs of courts. Note, however that, whereas one can characterize analytically, with relative ease, litigation cases with "external" benefits to society, it is by far more difficult to identify concrete "real-life" applications. For example, "deterrence of third parties" through law enforcement (embodying a positive external effect to the public) can only be achieved if the potential wrongdoers exhibit a sufficient degree of responsiveness, and this is an empirical question.

There may also exists an inverse constellation, in which the individual does not experience the full costs of her litigation, such that social costs are larger than social benefits. These incentives will result in "excessive litigation." Purely distributional conflicts, such as hereditary litigation, may serve as an example.

A special case of "external" benefits from individual litigation may arise if the defendant's action has not only affected one single victim, but also various other individuals (as is the case with many environmental damages), so that each individual victim's choice to enforce the law would generate benefits also for her peers. It is because of this "public good" aspect that the single individual may not be willing to provide law enforcements "for everybody" on her own costs.

Second, it is argued that, even with limited public subsidy, enforcement is costly, such that remedies measured only by the value of the property right are insufficient to induce adequate enforcement. For example, if one has a property right worth 10 €, but it costs 5 € to obtain legal vindication, there will be under-enforcement. In these cases, law enforcement will not be cost justified on the margin. Note that under this argument (meritorious) small claims are likely to remain unenforced irrespective of a possible divergence of private and social incentives for litigation. Even if the owner-claimant fully "internalizes" the enforcement incentive, he will rational refrain from law enforcement if this enforcement "eats up" the entire claim;

this analysis is only exacerbated in case of external social benefits of litigation. This argument often is used to justifying various forms of fee-shifting rules (whereby the losing party has to indemnify the prevailing party) in the procedural system (examined in Sect. 3.1).

Third, shortfalls in private law enforcement of whatever source entail consequences on the behavior of potential trespassers to be deterred by litigation, as potential defendants would weigh the amount of compensation due with the probability that they are successfully sued. Therefore, it is argued that, even with limited public subsidy and fee-shifting, imperfect enforcement may justify damages that are above compensatory levels in order to produce optimal deterrence of wrongful acts. For example, if enforcement is shown to be taken in only 50 % of cases, this is said to justify awarding damages at double the compensatory levels, in order to present the potential wrongdoer with an "expected liability" equal to the extent of the damages. Of course, if such "punitive damages" are raised above expected harm, this feature produces too much enforcement, which is equally as undesirable as too little enforcement.

A combination of all three arguments is sometimes used to justify the "class action" device used in the United States to aggregate large numbers of similar claims into a single adjudication. In practice, the procedure has a somewhat coercive aspect, in that absent claimants generally are required to "opt-out" of the class, which is costly to them. This is argued to be justified as eliminating a "free-rider effect" that otherwise would produce under-enforcement as the individual claimants waited for each other to commence the case, and therefore reduce enforcement costs to the follow-on claimants. The class action devices also introduces agency costs by allowing the few named claimants' lawyers to conduct the case on the part of the entire class, and to be paid first from the class recovery, often leaving little or no net recovery to the nonparty class members, and thus producing something akin to a "bounty" effect.

2.2 Public Enforcement

(a) Both the qualifications noted above and the counterarguments that they engender also surround the economic analysis of public enforcement, including enforcement by administrative action, by criminal prosecution, or by public civil action (whereby public agencies enforce the law "as a private party" under civil procedure which is more developed in Anglo-American than in continental European countries).

(b) Like class actions, public enforcement is said to be justified by economies of scale in enforcement costs (i.e., an additional unit of law enforcement can be more cheaply provided by an increase in output of one agency than by a separate supplier) and by the external social benefit of enforcement. Two other arguments often encountered in support of public enforcement are (i) limitations on the efficacy of private civil remedies, due to insolvency of the

defendant or the like, and (ii) the absence or ambiguity of assignment of private property rights to support private enforcement, as is in certain forms of environmental cases involving pollution to navigable waters or public lands.[2]

(c) However, somewhat like class actions, public enforcement has been criticized as introducing new problems that could produce either too much or too little enforcement. In these instances, the use of public agents (usually state officials) for purposes of law enforcement in the interest of the public (embodying the principal) introduces problems of agency cost and public choice. "Agency cost" refers to the problem that public agents' incentives may deviate from social welfare-maximizing incentives, and "public choice" influences can exacerbate such problems by selectively favoring certain types or times of law enforcement. Unlike private enforcement, public enforcement generally does not internalize either the benefit or the cost of enforcement to the enforcer. Public enforcement agents may be constrained by budgets to under-enforce, or by political or careerist pressures to over-enforce. These enforcement choices are, almost by necessity, surrounded with a certain side taste of arbitrariness (say, in case the police tickets only a few, but not all drivers for speeding or for parking violations). Furthermore, public enforcement priorities may reflect public-choice influences toward transferring the cost of enforcement from those obtaining concentrated private benefits of enforcement to the tax-paying population in general, and thus producing another form of wealth-transfer legislation (for example, lobbying by a certain city district for preferential police protection).

(d) All of these features of public enforcement also enter the analysis of criminal law enforcement by imposing punishment. If the crime already has taken place, then it can be argued that the crime victim's incentives may produce either under-enforcement or over-enforcement. Many people, at first blush, would expect over-enforcement of crimes, if punishment is left to the individual victim on the basis of an assumed vengefulness on her side. Some such motivation is likely to be present in the context of individually provided enforcement of the criminal law and may explain a certain amount of individual contributions to the enforcement of criminal law. However, this analysis disregard that punishment is costly (involving both various categories of "out of pocket costs" and intrinsic costs in terms of the "disutility" experienced when actually meting out punishment against the wrongdoer) and that the victim could save these costs if she leaves the wrongdoer unchallenged. In particular, if punishment does not produce a transfer for payment to the victim (perhaps because the criminal has no transferable assets or is subject to punishment by incarceration) it may be rational, at least, judged in

[2]The environmental pollution context provides one example of potential substitutability between substantive and procedural law, as one alternative to public enforcement could be the assignment of private property rights to the public good involved; these private property rights would, however, also require enforcement.

a mere response setting, not to pursue the wrongdoer.[3] Whereas the above-mentioned vengefulness may provide some incentives for law enforcement also in cases where the punishment, considered in itself, is not cost justified, the dominance of either incentive is an empirical question and not analytically predetermined ex ante. Moreover, even if the victim may explicitly wish the wrongdoer to be punished, she may prefer, due to the costs of punishment that someone else carry out the punishment so that she would not have to enforce the law herself. The relevance of the costs of punishment for actual choices can also explain why people on the street, when asked about their attitudes on punishing in a particular case, tend to be more vengeful (because of the lack of punishment costs) than as jurors on the bench (where they bear the intrinsic disutility of punishment).

For these several reasons, most procedural systems place criminal law enforcement in the hands of public agents. However, those agents still present the same problems of agency cost as in the more general case of public enforcement. Criminal law enforcement may be misused by the prosecuting agencies for political reasons against innocent individuals, as was the case frequently in earlier times, and still is, at some occasions, today. The current discussion in many continental countries is concerned with the inverse constellation, namely with possible shortfalls in the prosecution of criminal wrongs (which may result both from political bias to protect "friends," or simply from shirking). Among those institutional instruments currently considered to cope with insufficient enforcement by the public agency is the initiation of criminal prosecutions by judicial order ("Rechtserzwingungsklage" = "Action to enforce criminal prosecution"). More generally, if the enforcement agency is politically accountable (say, if prosecuting agents are directly or indirectly elected, as is the case in some parts of the U.S.), the election pattern is likely to influence enforcement activities (namely with the goal of shifting more resources toward the period immediately preceding the elections) and, ultimately, crime rates—the result being a politically influenced cycling of the crime rate.

2.3 Settlement Under the "Expected Value" Model

The most highly developed feature of the economic analysis of procedure is the trial-versus-settlement decision. The economic models take into account that

[3]More generally, if punishment is only seen in a mere response setting without regard to its deterrent effects, the disutility of punishment, as experienced by the punisher, is likely to dominate possible retributive concerns. If aware of this mechanism, the trespasser can even successfully exploit the punisher. This "punishment dilemma" helps explaining both the disutility of a mother educating her misbehaving child ("Wait until Daddy comes back, he will spank you!") and the seemingly puzzling empirical fact that jurors (lay judges) in some jurisdictions have been found to sentence more leniently than their professional peers on the bench, because of a systematic disregard for the deterrent effects of punishment.

litigation is a strategically interactive process, whereby each side reacts to the moves of the other. The trial-versus-settlement model attempts to identify the conditions under which some cases settle before decision, whereas others go to final decision.

2.3.1 Settlement in Private Civil Litigation

(a) The basic model of trial-versus-settlement postulates two opposing views of expected value (for the claimant) and expected loss (for the defendant). In each case, the "expected" result is a compound of each side's assessment of the value of the claim times the probability of a plaintiff victory at trial, as adjusted by each side's costs (which are analytically the costs of proceeding to the next stage of decision). In this basic model, the prospect of settlement is influenced by whether there is a "bargaining range" of overlap between the parties' estimates. Specifically, if the defendant's expected loss is greater than the plaintiff's expected gain, then the case will settle between those points of the bargaining range, unless prevented from doing so by either strategic behaviors on the part of the parties or external influences. On the other hand, in the absence of a bargaining range, then the parties can not settle without taking additional moves to change each other's estimates.[4]

(b) Much of the literature in the economic analysis of procedure is concerned with characterizing the conditions that prevent settlement. Three (mutually complementary) basic models have been proposed: (i) "prediction failure," (ii) "bargaining failure," and (iii) the "external effects" model. Whereas under the first two models there exists a bargaining range, but the parties fail to settle because of mutually inconsistent, relatively optimistic estimates as to the outcome of the trial or because of mere distributional quarrels, the third model argues that it is the existence of asymmetric external effects, and thus, asymmetric stakes that prevents the existence of a bargaining range and forecloses settlement.

"Prediction failure" refers to the case where the parties' estimates of the likely outcome do not converge sufficiently to produce a bargaining range. An important variation of this type is the so-called "mutual optimism" model, where each side believes that it is more likely to prevail at trial. The possibility of such a failure is one justification for the American system of "pretrial discovery" of evidence from the opposing litigant or third parties, which may help to dissuade one or both sides

[4]The analysis is further complicated by the existence of lawyers representing their clients. This legal representation can be explained as a "principal-agent-relationship" and there is much law-and-economics literature on this. The client-lawyer-relationship is one of asymmetric information regarding the lawyer's quality and costly monitoring that can produce either over-provision or under-provision of services. There are various methods, such as success-based remuneration or reputational markets that can produce convergence between the interests of both parties.

from their mutual optimism, but of course such a procedure increases litigation costs and opens up the possibility for strategic behavior within the discovery process, either by imposing asymmetrical discovery costs on one or another party or through selective disclosure of information.

"Bargaining failure" refers to the case where the parties' estimates converge or overlap but the parties nonetheless fail to settle because of strategic behaviors. To take the simplest case, suppose the parties actually agree with each other on the expected value and loss of the case. However, both parties still face incremental costs of proceeding to trial, and in this instance the true "bargain" is over how to allocate the mutual benefit of avoiding trial, which may or may not be symmetrical as between the parties. If one or another party bargains too hard over the division of that cost savings, or if one party is more or less averse to risk than the other, then there can be a bargaining failure. Or, the parties may behave strategically by failing to disclose their true estimate of the probable trial outcome, or their incremental costs of proceeding to trial. All of these can produce "bargaining failure."

A third "external effects" model postulates that important classes of cases may have "external" effects (meaning here effects materializing outside of the concrete litigation on the immediate parties to the dispute) that produce asymmetrical stakes and thereby eliminate a bargaining range even when the parties' estimates converge. One example where this can happen is the "repeat play" litigant opposing a "single play" litigant. For any of several reasons, such as the precedential or preclusive effect of an adverse or favorable judgment, or to establish a "tough" reputation to deter future litigation costs, the "repeat play" party may have more at stake than the "single play" litigant. Depending upon the relative magnitudes of the internal stakes versus external effects, such a situation can eliminate the bargaining range. One example might be where a bank litigates the validity of some provision in its standard loan agreement against one of its customers: the customer's stake is limited to the case at hand, whereas the bank may have an entire line of business involving thousands of customers at stake.[5]

[5]There has been a good deal of attention to the effect of "fee-shifting" rules (i.e., awarding litigation costs depending upon the outcome of trial) on the trial-versus-settlement decision, with ambiguous results. For example, a "loser pays" rule (as opposed to the "American rule" where both sides bear their own costs except in extreme cases) may do nothing more than raise the stakes for both parties, which may eliminate a bargaining range that might otherwise exist, or it may create an asymmetry of stakes if the two sides face differing cost functions or have differing attitudes toward risk, which may either discourage or encourage settlement.

2.3.2 Settlement in Public and Criminal Litigation

(a) Public litigation in general appears to be inherently more difficult to settle because most public litigation institutionalizes the asymmetry of stakes noted in the previous section through the intervention of the public enforcement agent. In the usual case of a public agency versus a private party, the public agent does not fully internalize either the benefit of a settlement or the cost of proceeding to trial, and therefore this type of litigation would appear to be influenced on the public side by an external effects model, with those external effects given by diffuse political or bureaucratic incentives.

(b) There has been extensive study of the U.S. criminal "plea bargaining" process as a special case of settlement of public litigation, though much of this literature suppresses the agency cost problems associated with the settlement of public litigation. In the basic model, the agency cost problem is addressed in part by assuming that criminal prosecutions are limited by finite prosecutorial budgets, and so the problem is cast as rationing these limited resources to a highly elastic supply of criminal offenders. In most basic models, the prosecutor is assumed to be maximizing overall deterrent effect by offering plea bargains as a sorting device to distinguish innocent from guilty defendants (or more guilty versus less guilty defendants). The guilty (or more guilty) defendants accept the plea bargain, which provides a discount from the expected punishment at trial, but this result is rationalized as allowing a larger number of guilty defendants to be convicted at a given fixed cost. The difficulty with this analysis is that, in order to make the strategy credible, the prosecutor actually would have to go to trial against innocent (or less guilty) defendants, and, if a conviction were obtained, would have to seek a high penalty, for otherwise the risk-preferring guilty would masquerade as the innocent. Both of those implications would appear to violate prosecutorial ethics (even in the U.S.) and justice considerations. Ultimately, the argument for such an analysis would seem to rest importantly on the assumptions that the trial process rarely or never produces either false positives (erroneous convictions) and that false negatives (erroneous acquittals) are relatively uncommon. Either or both of these assumptions appear to be debatable empirically. Because of concerns about the asymmetric costs of error, criminal procedure systems tend to arranged so as to tolerate a relatively high level of false negatives in order to minimize false positives.

Another view of criminal plea bargaining may place it closer to the standard model of settlement in private civil litigation, under the assumption that prosecuting authorities internalize a cost (perhaps to professional reputation) of taking the innocent (or less guilty) to trial or failing to convict the guilty (or more guilty), and therefore choose to plea-bargain not for sorting or signaling purposes, but simply on the basis of predictions of trial outcome, as constrained by limited prosecutorial resources.

3 Economic Analysis of Civil and Criminal Procedural Rules

(a) Virtually all civil and criminal procedural codes announce their objectives as securing the "just, speedy, and inexpensive" determination of legal disputes. These goals are interrelated. "Speedy adjudication" may also be inexpensive. And both speedy and inexpensive adjudication are regularly seen as aspects of "justice" in procedure, at least in the legal sense. However, an application of economic analysis demonstrates that these desiderata not only conflict with each other in some cases, but also are far more subtle than may appear at first blush. If "just" adjudication is seen in terms of reducing error costs, this reduction in error costs involves a tradeoff against reduction in direct costs of adjudication, which are more closely associated with the desiderata of "inexpensive" and "speedy" procedure.

(b) Closer observation shows that, as noted in the Introduction of this chapter, procedural systems and their constituent rules operate in conjunction with underlying substantive standards, and are characterized by interdependent strategic behavior by the parties and, to some extent, by the tribunals themselves. The basic goal of the procedural system is to make the substantive law efficiently enforceable, which does not imply the minimization of direct litigation costs only, nor even the sum of direct litigation costs plus "error" costs solely in the sense of erroneous factual or legal determinations in adjudicated cases. In addition, the procedural system "feeds back" on substantive law and influences both the supply of cases to the system and the supply of information available to resolve the cases that appear. Such a system would not be efficient, no matter how inexpensive and accurate it became, if, for example, it encouraged the bringing of cases that could have been more efficiently solved (or prevented entirely) by ex ante bargaining between the parties. Thus, a more inclusive statement of the optimization problem would be to minimize the total sum of transaction costs and other opportunity costs imposed by the legal system overall, both substantive and procedural taken together. At that level of generality, there is no completely satisfactory synthesis in the existing literature, because the problem is enormously complex. We can only look at pieces of the overall picture. In this section, we will examine some of the problems from the law-and-economics perspective, organized around the competing desiderata.

3.1 "Inexpensive" Adjudication

(a) As noted above, the most "inexpensive" adjudication is the one that never arises in the first instance. However, this does not imply an extreme solution. It may be more expensive for contracting parties, for example, to anticipate each

and every potential dispute that may arise, most of which will never arise. But the parties' incentive to anticipate ex ante is given in part by the cost of ex post litigation. If ex post litigation were very cheap, then the parties would have little incentive to anticipate, and there would be a great of litigation, much of it concerned with trivial matters that are unworthy of anyone's attention, including the immediate parties. If it were too expensive, then parties may not contract at all, and may not engage in any number of other social interactions. Obviously, the cost of litigation is not the only influence: clearer substantive law also tends to reduce both the supply of disputes and the cost of their resolution, perhaps more dramatically than any procedural rule.

(b) To the extent that the costs and benefits of litigation were entirely internalized to the immediate parties, and distributed in such a way as to minimize strategic behaviors,[6] then social policy might leave the entire problem to the parties' private agreement and need not bother about the costliness of the procedure. This would be the purely "Coasian" solution. In the absence of some external effect on a third party or a defective bargaining process, there seems to be little basis for compelling private disputes to be litigated in public courts. This implies a policy that is open to contractual substitutes for litigation, such as arbitration, mediation, or the like.[7] In fact, such a policy is followed by most procedural systems.[8]

It would further appear that agreements between the parties, either before or after the litigation, generally should be respected by the public court system, no matter how "expensive" they appear, unless they also impose undue public expense. In other words, parties should be permitted to contract for their own procedure, within limits. This stands in contrast to the case of one contesting party's unilateral request for expensive procedures, opposed by the other party. This instance may present a case of strategic or opportunistic behavior, but not necessary by the party who requests the "expensive" procedure. By analogy to the economic analysis of tort or contract law, the appropriate approach would seem to be a reconstruction of what the parties would have agreed to before the

[6]However, there are cases (such as in tort law) where high ex ante transaction costs prevent that solution, and there are cases in which an external benefit (through formulation or clarification of legal rules for the benefit of third parties) perhaps would be lost or under-provided.

[7]In arbitration, the litigants regularly opt for a different institutional mix regarding the goal of minimizing total costs. Whereas in state courts, there is a division of labor (and a split of costs) between trial courts and courts of appeal, the parties of arbitral proceedings tend to divert the resources of the appellate level to a more extended procedure in the first instance where the parties usually submit to a panel of three experienced arbitrators, one of whom each party assigns, the chairman being determined by the two other arbitrators.

[8]A very substantial proportion of national and international commercial litigation is decided by arbitral tribunals. One factor that contributed to the success of arbitration is the near-universal recognition of their awards under the New York Convention. Arbitral tribunals are both ad hoc tribunals (contractually agreed upon but established only at the occasion of the dispute) or institutionalized arbitration that provide a set of general procedural rules and a "hosting" institution that sponsors the selection of the panel to decide the concrete dispute.

dispute arose, which may be somewhat easier to determine in a procedural than in a substantive context (see Sect. 3.3.2). Failing that approach, proportionality to the size of the stakes in the dispute is a rough guide to identifying excessive expenditures.

(c) Setting aside the caveat noted above regarding "inexpensiveness" as a value in itself, one can discuss the legal desideratum of "inexpensive" litigation along different lines.
In a first, quite immediate sense, this desideratum suggests the provision of a "lean" procedure, which does not only economize on those elements and costs "unnecessary" for a specific type of litigation, but also provides for the creation of different categories of courts (from "small claims" courts to "Supreme" or "Superior" courts), with procedures that vary in complexity and cost with the amount in controversy. Here again, some measure may be found in comparing litigation costs with the disputed stakes. For the economist such types of courts embody a distinct trade-off between direct and indirect costs of litigation. The above differentiation of various categories of courts is also of assistance in allowing (meritorious) low value claims to be litigated.

(d) In the existing literature, the "inexpensiveness" of litigation is often interpreted as a question of equal access to the courts, and there is concern about raising the minimum cost too high, out of fear of pricing the non-affluent out of the courts. It is possible to develop an economic argument for some such institutional device in the light of the otherwise pending erosion of property rights. If only the affluent could litigate their claims, substantive property rights of the non-affluent would systematically be undermined. Still, it is a difficult institutional question how to ease this type of litigation. Different jurisdictions take different approaches to the problem. Under most European systems access to the legal system is granted by a complex system of either direct public provision of legal services to lower income individuals (at least in certain legally more demanding or high-valued cases) or by pro bono activities on the side of the legal profession. Regularly, these systems also rely on fee-shifting rules. In the United States, public legal aid is much more limited and legal pro bono activities largely are left to the selection of individual members of the legal profession. In the U.S., the contingent fee system[9] (mostly encountered in torts cases) seems to function as a financing vehicle that sorts cases for threshold merit and spreads risk by inducing plaintiffs' lawyers to build a diversified portfolio of cases. However, some jurisdictions are hostile to the contingent fee, including most of Europe. Most jurisdictions permit similar arrangements on the defense side by liability insurance, which is expressly designed to spread litigation risk.

[9]Under this system the attorney is paid according to success. In the standard conditional fee contract the attorney's reward in case of success amounts to a certain percentage to the claim (mostly around 33 %).

(e) In a further important sense the criterion of "inexpensive" litigation requires that the costs of the legal system incurred for the enforcement of a legal claim should not be prohibitive. One instrument to accomplish this goal is the implementation of a "fee-shifting-rule." Under such a rule, the winning litigant is entitled to recover her attorney's fees and out of pocket costs (including court fees). Under the "American rule,"[10] in turn, there is no such indemnification so that each party bears her own costs.[11]

The consequences of fee-shifting are to some extent ambiguous.

First, fee shifting cannot be seen independently of the substantive law. Its effect on case supply seems to be toward confirming the legal status quo, which itself is ambiguous. One could argue that fee-shifting has favorable efficiency properties for the enforcement of established law, to the extent that it compensates for the cost of enforcement that may otherwise reduce incentives for private enforcement below optimal levels. However, this assessment depends in part upon the correspondence between legal damages and economic damages: if the legal damages formula under-compensates, then fee-shifting may ameliorate but does not correct the problem; on the other hand, if the legal damage formula over-compensates, then fee-shifting makes the over-enforcement problem worse. It also depends upon how closely established substantive law corresponds with the economically efficient rule, as it tends to make established law more resistant to change.

Such ambiguities also exist with respect to the procedural consequences of fee shifting in a narrower sense. The "loser pays" fee-shifting system may tend to discourage settlement by raising stakes, and tends to increases the returns to more certain claims relative to other claims, especially for more risk-averse (usually meaning lower income) claimants, who otherwise are selectively more discouraged by fee-shifting. In addition, fee-shifting system have some tendency to encourage what economists call the "moral hazard"(also characteristic of some insurance systems), which means that individuals are likely to decide to spend more if they believe that someone else ultimately will have to pay. To the extent that one side's litigation expenditures also induce the opposing party to raise its expenditures, this

[10]It should be noted in passing that the "American rule" as encountered in the literature is only an approximation of the actual practice in American courts. While the general rule in America is that each party bears its own costs, this is subject to several exceptions, most notably the "bad faith" exception, which seeks to screen out dishonest or ill-founded claims for fee-shifting treatment. Similar provisions are found in most American procedural codes by provisions for fee-shifting or other sanctions upon both lawyers and parties asserting "frivolous" claims or defenses. Finally, in a number of areas, legislation has been enacted to permit "one-way" fee-shifting in favor of plaintiffs successfully asserting certain specified types of claims (for example, antitrust claims, civil rights claims) that are thought to be under-provided by the usual incentives of private enforcement.

[11]The problem of under-compensation for enforcement costs under the American rule has been one argument made in favor of the more common practice in the United States of awarding "punitive" (i.e., higher than compensatory) damages in tort case, or what is known as the "collateral source rule" (which does not offset tort damages for insurance reimbursement or the like). However, neither measure seems well-adjusted to the problem.

effect could be magnified by use of litigation expenditures as a strategic weapon. For these reasons, most fee-shifting systems embody a legal limitation to "reasonable" expenditures (regularly based on lawyers' tariffs). Although the limitations provided in most European systems work smoothly, one must admit that the criterion of "reasonable" expenditures has sparked many legal disputes on cost issues, raising the direct costs of administering the system.

Still, there are more straightforward aspects of fee shifting. Due to the obligation to indemnify the prevailing party, fee-shifting discourages potential litigants from filing non-meritorious claims (meaning claims with low probability of winning). In turn, fee-shifting encourages meritorious claims, because it grants the plaintiff the prospect of being fully compensated for his costs (to the extent that the officially recognized tariffs represent the true market value of the services supplied) and to receive the reward. Therefore, economic analysis indicates that the European alternatives to the American contingent fee plus the "American rule" (under which parties bear their own costs) may selectively favor more certain claims over more speculative claims. This distinction may be justified by the differing approaches of the two systems toward case decisions as primary sources of law: in the U.S., most tort law is "common law", i.e., based entirely on case decisions rather than statutes; in such a system, it may be necessary to encourage a wider diversity of cases in the first instance, in order to provide more raw material for the formulation of legal doctrines. One drawback to the European systems could be the higher direct costs of administration: public assistance cases involve public bureaucratic and political costs of budgeting, screening, and organization; fee-shifting systems involved more direct cost of judicial administration. The American contingent fee is a private market system that involves little or no direct public administration or supervision.

3.2 "Speedy" Adjudication

(a) The desideratum that procedure should be "expedient" is close to universal recognition in legal doctrine both in the realm of civil and criminal law. Most legal systems even have a saying that "justice delayed is justice denied." Still, on a more abstract level, the criterion of "speedy adjudication" is more ambiguous. Litigation can be "inexpedient" because one side deliberately delays the procedure or because both sides prefer a slower pace of their litigation. At least under European continental procedural systems, the inexpedience of a trial can also be the result of mismanagement of the case on the side of the judge, be it that he fails to implement an effective schedule of the proceedings (or is himself an obstacle because of bad preparation or illness), or

fails to enforce the procedural instruments available to cope with unilateral attempts from one side for undue delay.

(b) Much of what has been said about "expense" also could be applied to the supposed objective of expeditious litigation. Here again, in private civil litigation, the main problem for procedural rules would be to distinguish between purely strategic delaying tactics and legitimate (or mutually agreed) development of a case. There is an obvious conflict between a "speedy" disposition and a "just" one. But to some extent, there also is a direct conflict between a "speedy" adjudication and an "inexpensive" one, if the desire for "speed" in itself forces the litigants to incur unnecessary litigation expense that could be avoided (for example, in cases where both sides may prefer a slower development, because other events may avoid the need for a definitive adjudication). Unless such mutually agreed delays affect third parties, as by depriving other litigants of access to public judicial resources, there does not seem to be any public interest in pushing private litigation along.[12] Moreover, while it seems perfectly acceptable that procedural systems should be arranged so as to minimize the use of dilatory tactics as a form of predation by one party against another, it is equal possible that expeditious tactics also can be used as a weapon of predation. Therefore, institutional rules also must cope with this (inverse) constellation.

(c) Criminal cases may appear to differ in this respect, given the emphasis on "speedy trial" guarantees in criminal procedure. However, in most such instances, this is a right in the accused against dilatory tactics by the prosecuting government. One might still argue, although somewhat attenuated from the economic perspective that there is some independent external interest of the general public in expeditious disposition of criminal charges, even in the absence of the defendant's objection or perhaps because of concerns that the defendant's consent may be extorted by the government in ways that cannot be directly observed.

(d) Currently, we observe increasing calls for more expeditious dispositions, both in the European system and in the United States. When assessing the quality of this reform, one has to be aware of the heterogeneous nature of "inexpedience" in proceedings outlined above. Insofar as this reform is directed against judicial ineffectiveness, the preservation of judicial independence limits the deployment of direct performance-dependent ("carrot and stick") incentives with respect to

[12]The importance of this factor may depend upon the nature of the procedural system and the stage of the case's development. In American procedural systems, where preliminary proceedings are conducted largely without judicial involvement and there is a sharp distinction between the "pretrial" and "trial" stages, there seems to be no case for placing the parties on a judicially mandated timetable prior to trial, unless one or both parties request such a schedule. In this respect, European procedure may differ, as there is less of a distinction between pretrial and trial, and more active judicial involvement throughout. Still, also in private litigation under the rule of continental European procedure, there is no clear public interest in prompt disposition against the wishes of both parties, unless their delay prevents another case from advancing in the queue.

judicial behavior.[13] One possible explanation for the recent reform may lie in the incentives of the judges and judicial bureaucracies themselves, as case dispositions and intervals to disposition are one of the few quantifiable "outputs" of the judicial system.[14]

3.3 *"Just" Adjudication*

(a) It is commonly recited that procedural codes should ensure a "just" outcome. While the general concept of what constitutes justice can be controversial in substantive terms, it is the case that aspects of justice in procedure command more common agreement. From the economic perspective, a "just" outcome is most closely connected with the idea of reducing error costs.

(b) Procedural justice generally has three aspects: (i) correct application of substantive law; (ii) correct determination of the objective facts (the "material truth") of the case; and (iii) procedural fairness in the formulation and application of the procedural rules themselves. Our discussion below focuses on the second and third aspects of procedural justice.

[13]The pertinent economic literature generally assumes that the behavior of judges can be explained precisely along the same lines as the behavior of ordinary people. From the economic viewpoint, judges maximize their utility (which encompasses several elements, such as income, promotion, prestige, avoidance of reversals, perhaps concern for fairness) under given constraints. These constraints are under most laws such that judges are immunized against direct performance-dependent incentives to secure their independence vis-à-vis political influences. Judicial compliance is secured by a system of more indirect incentives, regularly relying on postponed remuneration (where generous pension arrangements make it unattractive to drop out of the judicial career due to some misbehavior) and on monitoring schemes, to which peers, senior officials, and appellate courts contribute. Since judges regularly do not have fixed working times, they are partially remunerated by leisure, which creates an imperfect incentive device for expedient working.

[14]If judicial evaluation, funding, or personnel is determined by some measure of "throughput," then judicial bureaucracies may have an incentive to make their dispositions more "speedy", even if by doing so they are socially more expensive and erroneous, as those consequences are not as fully internalized to the judiciary. An alternative explanation could be that certain courts are attempting to attract certain types of judicial business by "signaling" to potential litigants or classes of litigants their willingness to accelerate either all cases or certain types of cases. Something like this effect may explain why the U.S. federal government has chosen to prosecute several of its recent terrorism cases in a certain federal district in Virginia (one of some 100 federal districts) that has cultivated the reputation of providing a "rocket docket," thus inviting certain classes of litigants who particularly value speed. This is one way that judges can effectively "select" the types of cases they would like to hear, where jurisdictional competency is non-exclusive, which is often the case in the United States.

3.3.1 Justice by "Material Truth"

(a) The concept of "material truth" is that there is an objective truth of the factual grounds for decision in a given case. It should be noted that this concept does not play an equally central role in all systems of procedure. The Anglo-American procedure is more concerned with determining the facts as they are submitted to the tribunal by one or both parties, especially in private civil cases. Within that system, it is permissible for the parties to stipulate to a set of facts, which are submitted for judicial decision. Nevertheless, the facts submitted to the tribunal generally are claimed to correspond with actual events, and the explicit decision of purely hypothetical cases is prohibited.[15]

(b) If we focus only on the fact-finding aspects of adjudication, there is some debate within the law-and-economics literature over the extent of the external social interest in minimizing factual error costs, particularly in private civil disputes. Most of the cost of such error falls on the immediate parties. Provided that the level of inaccuracy in adjudicative fact-finding is not so extreme as to virtually force parties to shift away from adjudication to substitute forms of dispute-resolution (such as vendettas or feuds, organized crime, and so on) with negative social consequences, there may be a broad range of tolerable accuracy levels. Presumably, to the extent that parties are permitted to contract away from public courts to peaceful alternatives such as arbitration, the accuracy levels of public adjudication are disciplined by competitive forces.[16] However, in areas of public monopoly in procedure, such as criminal cases, it may be more important for the procedural system to embody internal controls on factual accuracy, and this is consistent with some of the features that distinguish criminal from civil procedure, as in the more extensive procedural protections for the accused designed to minimize false positive errors (erroneous convictions).

(c) Moreover, in all systems, the "search for truth" is limited by competing criteria of cost and efficiency. Most systems focus factual inquiries, whether by judges or parties, on those facts identified as pertinent or relevant by the applicable substantive law. Thus, parties may wish to show, outside of the narrow "pertinent facts of the case" that they are particularly worthy or virtuous individuals, but generally speaking, such considerations are excluded by substantive law and, therefore, not the object of proof in procedure. In "common law" systems

[15]This is conceived as a limitation on the jurisdictional competency of courts, which could be justified on any of several grounds: as rationing access to public decisional resources, ensuring adequate incentives to the parties, or protecting the reputation of courts as reliable dispute-resolution institutions.

[16]Institutional competition exists also among institutional arbitrational tribunals (such as those established at the Chambers of Commerce in leading European capitals). Since the hosting institutions (the Chambers) derive a direct benefit from litigation in terms of court fees (there are no public subsidies for arbitration) and parties go after what they deem the most efficient procedure, competition for litigation among these arbitrational tribunals has contributed to a remarkable convergence of the respective arbitrational codes.

where case decisions are conceived as one source of law production, the criterion of relevancy may be more loosely applied, so as to permit the evolution of legal rules through case law. This feature may be one of those contributing to a higher direct cost of adjudication in such systems, but it is not clear whether these costs compare favorably or unfavorably to the costs of fact-gathering for the purpose of generating legal rules by legislation.

(d) Further, in all systems, the "search for truth" is not self-executing: that objective must be implemented through the design of mechanisms within the procedural rules themselves for generating factual information and resolving disputes. In particular, the procedural code influences the incentives for all persons involved (including the tribunal and witnesses, as well as the parties) to contribute toward that end. Obviously, the parties are self-interested and opposed, and can be expected to engage in strategic behaviors and selective disclosure of evidence. However, the same may be true of nonparty witnesses, who may have interests pertaining to the case or to their testimony. Nor are individuals associated with the tribunal itself immune to incentives: both professional judges and lay judges (jurors) may have incentives to deviate from "truth," depending upon the institutional structure surrounding their work. For example, if judges were evaluated and promoted solely on the basis of their "throughput" of cases, then judges' incentives would be more strongly aligned toward solving the case quickly rather than accurately. On the other hand, it has been argued that judges with little or no prospect of promotion, secure tenure, and minimal case-processing requirements may have an incentive to work as little as possible, and when they work, to spend more time on "interesting" cases, which also does not seem favorable to achieving high levels of accuracy.

(e) Procedural systems differ markedly in their assignment of roles and incentives to the various actors in the fact-finding process. In particular, there is a contrast between more "adversarial" procedures versus more "inquisitorial" procedures. While most actual procedural systems embody some mix of these two types, traditionally the Anglo-American system emphasized "adversarial" elements while the European continental system emphasized "inquisitorial" elements. The adversarial system relies primarily upon the competitive interests of the opposed parties to bring out factual information, while the inquisitorial system places heavier reliance on the neutrality, professional training, and questioning rights of the adjudicating judge to ensure a full revelation of "material truth."

(f) The choice between these two types of procedural approaches appears to involve a complex set of tradeoffs between direct and error costs of adjudication, and there is very little empirical knowledge on the subject. Theoretical literature suggests that inquisitorial procedures may reduce direct costs while maintaining a tolerable but not maximal degree of accuracy. On the other hand, adversarial systems, though probably more expensive in terms of direct costs, may achieve higher factual accuracy under certain conditions, though it is not clear that such a tradeoff would be economically efficient. One difference to be observed between

the systems is the difference in attention given to fact-finding in the original instance versus appellate review: inquisitorial systems involve more extensive appellate review of fact-finding, whereas adversarial systems tend to focus appellate review on questions of law only.

(g) The available empirical evidence is very limited, but is consistent with the view that, leaving aside the cost aspects of litigation, the relative efficiency of adversarial versus inquisitorial procedures varies with surrounding conditions. This evidence consists of experimental research in which "material truth" could be supplied artificially.[17] The method of investigation involved simple case scenarios with hidden facts known only to one of two opposing parties, with the roles of both parties and "referee" played by experimental subjects (mostly University students) according to stylized rules of purely "adversarial" (completely passive referee; questioning by parties only) versus "inquisitorial" (active questioning referee, but no questioning rights in parties) procedure. Referees were then required to award all or part of a contested stake to one or both parties, in accordance with a simple given rule of law. Under the full revelation of "material truth," the accurate decision was to award the entire stake to one party, and nothing to the other, under the given rule of law. While the experimental subjects concentrated on their private incentives (parties to "win" the monetary stake; referees to render the "accurate" decision, which maximized their payoff), the experimenters also observed the incidence in which revelation of the important hidden fact occurred under the respective questioning systems. The results were that relative revelation rates depended upon the degree of information asymmetry with which the parties began: where the hidden fact was exclusively on one side, without any hint to the other, then inquisitorial procedure achieved a higher rate of revelation; however, when the less-informed party started with a slight "clue" to pursue, then adversarial procedure achieved a higher rate of revelation, both relatively and absolutely. One institutional interpretation of these results is that adversarial procedures perform much better when preceded by the opportunity for the type of pretrial "discovery" that is characteristic of American civil procedure. However, this aspect of American procedure is widely believed to raise the direct costs of adjudication, and may also tend to suppress the ex ante production of information prior to the dispute. Therefore, once again the efficiency properties are ambiguous.[18]

[17]For reports of these findings, see Block et al. (2000), Block and Parker (2004), Parker and Lewisch (1998).

[18]Another interesting finding from this experimental research was that when revelation was achieved, then both systems tended to obtain roughly the same level of accuracy in the experimental referees' decisions. When revelation was not achieved, both systems had roughly the same level of inaccuracy, but the errors were distributed in slightly different ways. In particular, the errors of adversarial decision tended more strongly toward a "split the difference" outcome. This finding suggests a more important role in adversarial systems for placement of a "burden of proof" on the plaintiff, which is the observed general rule in Anglo-American systems. Without such a rule, adversarial systems may unduly encourage the bringing of weak cases simply to obtain a "compromise" verdict.

With the respect to evidence law, it is characteristic for many European countries that the respective legal rules are not uniform, but enshrined in the separate procedural acts (say, in an act on civil procedure, on criminal procedure, and on administrative procedure). In contrast, the Anglo-American system includes a highly developed body of evidence law doctrine that regulates the forms of evidence that may be presented by the adversarial parties. This body of law applies both to civil and criminal procedure, though it contains some special rules applicable to each type. Aside from the standard of relevancy, this body of law focuses extensively on the acceptable forms of evidence, preferring extemporaneous live testimony based upon witnesses' first-hand sense impressions to any other form, such as written testimony or deposition, documentary evidence, or opinion testimony by experts. To some extent, the development of this body of procedural law may have been influenced by the extensive use of lay juries as the fact-finding institution in Anglo-American procedure, which necessitates a compact "trial" stage in order to minimize the burden of jury service on citizens. However, it also reflects the parties' primary responsibility for evidence production and presentation in the adversarial system, through the incentives provided to the parties. Thus, the famous "hearsay" rule in Anglo-American law operates not so much as a rule of exclusion of "hearsay" (out-of-court) statements, but rather as an incentive for the party who seeks to benefit from the statement to bear the initiative and costs of actually producing the witness in court, which also permits the opposing party to subject the evidence to adversarial testing by cross-examination.

3.3.2 Justice as Procedural Fairness

(a) Notwithstanding the differences across procedural systems in the mechanisms used to achieve fact-finding, there is a remarkable degree of commonality across systems in the content of "procedural fairness," such that most systems embody the idea of symmetrical access to procedural rights between opposing civil litigants, limitations on favoritism to certain classes of litigants, and recognition of the need to provide expanded rights to the criminal accused. This feature is in some contrast to divergent notions of substantive "fairness" that can be observed in many areas of substantive law. Economic analysis can help to explain why procedural fairness has a wider degree of universality.

(b) The fairness or equality properties of legal rules are likely to suffer when the distributional consequences of such rules are transparent upon enactment, as in the case, for example, of tax laws. Both interest groups and decision-makers can easily predict the unidirectional wealth-transfer effects of such rules, and

this accounts for much modern legislation of this type. However, the more general and multidirectional is the prospective applicability of the rule in question, the more likely it is to be fair. Thus, for example, general rules of contract law are more likely to be fair and neutral than rules of tax law or regulatory law, because virtually anyone could be on either side of a given issue of general contract law.

There is an analogy here to the idea of "justice as fairness" put forth in the political philosophy of John Rawls, who introduces the concept of the "veil of ignorance." According to this argument, principles of justice in the arrangement of a society will be promoted by deciding behind a "veil of ignorance" that prevents each individual from knowing their own personal characteristics (e.g., social standing, age, sex, income, health, intelligence, and so on). In that instance, even self-interested choices of social arrangements will be unbiased to personal attributes and therefore "fair."

Constitutional economics has extended this concept to characterize a setting of rule-choice that mimics the properties of the "veil of ignorance," by postulating two conditions: (i) an extended time dimension of the rule (its longevity); and (ii) the generality of the rule's application. In combination, these conditions impede the decision-maker's ability to forecast the rule's distributional consequences for given individuals.

(c) Now let us apply these ideas to procedural rules. If such rules are long-lived and general in application, as they tend to be, then their formulation is more likely to reflect an idealized concept of justice as fairness. In the case of civil procedural rules, a given individual (including the rule-maker) would be unable to predict whether he or she ultimately would be a plaintiff or defendant, and therefore would be less like to support or promulgate rules systematically favoring one type of party over another. In this context, arguably it is in everyone's self-interest to agree on procedural rules that are fair in general, though it does not exclude the possibility of special-interest rules in certain predictable classes of special-interest cases.

(d) We can now compare this principle against observed instances of general procedural rules. One example is provided by Article 6 of the European Convention on Human Rights, which embodies, inter alia, the principles that no one should be the judge of her or his own cause, and that both parties to a dispute should be heard before a final judgment (audiatur et altera pars). We observe similar basic norms in most procedural systems. These are completely self-interested provisions, from the standpoint of someone who does not know the role they may play in relation to the procedural system.

More peripheral procedural provisions (for example, the right of appeal, or the exact timing of hearings) may be more debatable, and may require the distinction of criminal from civil procedure.

Many provisions in criminal procedure focus on the goal of minimizing false positive errors (i.e., convicting the non-guilty), which reflects some asymmetry of procedural rights between the prosecution and the defendant. Under an

extended time frame and generality, plus the severe and asymmetrical costs of error, these provisions may coincide with self-interest. Many individuals may believe that they are relatively unlikely to be criminal defendants. However, under a highly extended time frame, the cumulative probability rises. Furthermore, given the focus of such provisions on protecting the innocent, a long-term view may evaluate the benefits of such protections as outweighing the costs in general. In this respect, one might contrast such rules as appellate review and a heightened standard of proof with a rule excluding improperly seized evidence, which does not differentially benefit the innocent or contribute transparently to improved accuracy, and may well have the opposite properties. Thus, whether or not such a rule is justified on balance as a matter of social policy (say, as influencing the incentives of police and prosecutors in more positive directions), we should not be surprised to see some diversity of approach across different procedural systems.

Nor should we be surprised to find some diversity in more detailed rules across different systems of civil procedure. For example, a particular system may devote more resources to fact-finding in the first instance, and thus limit or dispense with appellate review on facts. These types of rules are more likely to be dependent upon local conditions and traditions, and in fact this is what we observe.[19]

(e) The distinction between generally shared notions of substantive versus procedural justice redirects our attention to the substitutability between substantive and procedural law, and helps to explain why many constitutional provisions are procedural in nature. Substantive constitutional provisions may be less common simply because it is easier to forecast their distributional consequences and therefore more difficult to agree on their content. Hence, in constitutions we observe a higher proportion of rules governing the production of substantive rules (for example, representational standards, required pedigrees for legislation, and so on), rather than substantive rules themselves. In this sense also, procedure substitutes for substance.

[19]For example, in the United States, the general pattern of civil procedure devotes extensive resources to fact-finding procedures in the first instance, devotes little attention to appellate review of facts, and defers appellate review until after the final judgment of the court of first instance. However, some important states, such as New York and California, deviate from this pattern. Both of those states freely allow "interlocutory" appellate review to interrupt the first instance proceedings, and New York allows one level of appellate review of fact-finding. These variations may reflect differences in either the procedural system or its surroundings, such as the types of personnel available to trial or appellate courts, or the nature of the cases supplied to these systems.

References

As basic references see the entries in the Encyclopedia of Law and Economics V for 'Civil Procedure General' (Kobayashi/Parker), 'Criminal Procedure' (Lewisch), and 'Evidence' (Parker/Kobayashi).

The more interested reader shall also consult the entries: 'Judicial Organisation and Administration' (Kornhauser), 'Appeal and Supreme Courts' (Kornhauser), 'Indemnity of Legal Fees' (Katz), 'Settlement' (Daughety), Arbitration' (Benson), and 'Class Actions—Representative Proceedings' (Silver).

See also Posner, Richard. 1998. *Economic Analysis of Law*, 5th ed, 563–672. Part VI 'The legal process'.

Further References

Block, M.K., J.S. Parker, O. Vyborna, and L. Dusek. 2000. An Experimental Comparison of Adversarial versus Inquisitorial Procedural Regimes. *American Law and Economics Review* 2: 170–194.

Block, M.K. and J.S. Parker. 2004. Decision Making in the Absence of Successful Fact Finding: Theory and Experimental Evidence on Adversarial versus Inquisitorial Systems of Adjudication. *International Review of Law and Economics* 24: 89–105.

Clemenz, G., and K. Gugler. 2000. Macroeconomic Development and Civil Litigation. *European Journal of Law and Economics* 9: 215–230.

Cooter, R., and D. Rubinfeld. 1989. Economic Analysis of Legal Disputes and their Resolution. *Journal of Economic Literature* 27: 1067–1097.

Landes, William. 1971. An Economic Analysis of the Courts. *Journal of Law and Economics* 14: 61–107.

Posner, Richard. 1973. An Economic Approach to Legal Procedure and Judicial Administration. *Journal of Legal Studies* 2: 399–458.

Shavell, Steven. 1982. The Social versus the Private Incentive to Bring Suit in a Costly Legal System. *Journal of Legal Studies* 11: 333–339.

Tullock, Gordon. 1980. *Trials on Trial: The Pure Theory of Legal Procedure*. New York: Columbia University Press.

In German

Adams, Michael. 1981. Ökonomische Analyse des Zivilprozesses.

Ministry of Justice, P. Lewisch and W. Rechberger, eds. 1998. '100 Jahre ZPO: Ökonomische Anlayse des Zivilprozesses' with the contributions by Clemenz/Gugler, 'Gesamtwirtschaftliche Entwicklung und Prozeßverhalten', and Parker/Lewisch, 'Materielle Wahrheitsfindung im Zivilprozeß'.

Misconceptions About Emissions Trading in Europe

Edwin Woerdman and Andries Nentjes

1 Introduction

In the previous century, law and economics scholars postulated that transactions of emission entitlements would improve the cost-effectiveness of environmental regulation. In their seminal work, Calabresi and Melamed (1972), building upon Coase (1960) paid considerable attention to the idea of creating pollution markets. On the basis of largely positive experiences with such markets in North America since the seventies, European policymakers picked up the idea in the nineties when they started preparing an emissions trading scheme for greenhouse gases to combat climate change. The result was the establishment of the European Union Emissions Trading Scheme (EU ETS), which became operational in 2005 (e.g. Faure and Peeters, 2008).

The original EU Emissions Trading Directive determines the rules applying in the periods 2005–2007 and 2008–2012 (Directive 2003/87/EC). The European Parliament and the Council of the European Union have since adopted an amending Emissions Trading Directive, introducing some new rules to apply in the period 2013–2020 and beyond (Directive 2009/29/EC). In 2015, the European Commission presented a legislative proposal to revise the EU ETS for the period 2020–2030 (COM (2015) 337), which has not yet been adopted at the time of writing this chapter.

Although the market for greenhouse gas emission rights in Europe is up and running for more than a decade, various academics, policymakers and observers argue that the EU ETS is not functioning as it should. There are indeed several

E. Woerdman (✉) · A. Nentjes
University of Groningen, Groningen, Netherlands
e-mail: e.woerdman@rug.nl

A. Nentjes
e-mail: a.nentjes@online.nl

A. Marciano and G.B. Ramello (eds.), *Law and Economics in Europe and the U.S.*,
The European Heritage in Economics and the Social Sciences,
DOI 10.1007/978-3-319-47471-7_12

implementation problems, but the debate is blurred, over-simplified and sometimes even misguided. In particular, there appears to be much confusion about what the EU ETS should deliver, whether it is effective in doing so, and how the scheme should be characterized. The aim of this chapter, therefore, is to confront and correct a number of these misconceptions about the EU ETS. An example is the idea that the EU ETS is a cap-and-trade scheme, which aims to incentivize low-carbon innovation. The chapter will explain, based on economic theory and legal rules, that this is not—or to be more precise: only partly—the case. By unravelling the complexities of the scheme, this chapter intends to contribute to a more nuanced and more accurate debate on emissions trading in Europe. We perform a positive legal-economic analysis (Backhaus 2005) and build upon earlier research by gathering, linking and fine-tuning insights from a number of studies on emissions trading (including Woerdman and Nentjes 2016; Woerdman 2015; Nentjes and Woerdman 2012).

This chapter is organized as follows. Each section title states a popular misconception about the EU ETS which we try to correct or nuance in the text that follows. Section 2 criticizes the first misconception that EU ETS is not effective as an instrument to restrict carbon emissions. Section 3 discusses the second misconception that the objective of the EU ETS is to stimulate low-carbon technological innovation. Section 4 reflects upon the third misconception that the EU ETS is a cap-and-trade scheme. Section 5 concludes.

2 First Misconception: 'The EU ETS Fails to Adequately Restrict Carbon Emissions'

Various authors and actors argue that the EU ETS has failed as a climate policy instrument. For instance

> The EU Emissions Trading Scheme has failed: "Time to scrap the ETS". (Lang 2013: 1).
>
> The ETS has long been a mess. Partly because recession has reduced industrial demand for the permits, and partly because the EU gave away too many allowances in the first place, there is massive overcapacity in the carbon market. (…). Prices had already fallen from € 20 ($30) a tonne in 2011 to € 5 a tonne in early 2013. (The Economist 2013: 1).
>
> (…) failing EU carbon market threatens effectiveness (…). (Greenpeace 2013: 1).
>
> The initial EU ETS has not been successful because too many allowances were distributed, and (…) a better response to global climate change would be a carbon tax (…). (Avi-Yonah and Uhlmann 2009: 49–50).

It would be unfair to throw all of these statements on one big pile and consider them equal, since they are not. The citations do not only come from a variety of sources (NGO's, journalists, academics), but they also differ substantially from each other: some are normative whereas others are positive, some are outspoken whereas others are more nuanced, some hint at repairing the EU ETS whereas others wish to

replace it with a carbon tax. What they all share, though, is that they claim some kind of failure of the EU ETS to adequately restrict carbon emissions. We will not so much argue that their views are wrong, but rather that they are imprecise, from a law and economics perspective. The aforementioned statements only tell half of the story, as will be demonstrated below.

The first decade of the European carbon market has been a period with a number of surprises. In Phase I (2005–2007) of the EU ETS, a learning period prior to the first commitment period of the Kyoto Protocol, the allowance price initially rose from € 10 in 2005 to € 30 in 2006, but the allowance price fell to less than 1 euro in September 2007 once it became known how generous allowance allocations had been. In the EU as a whole, the number of allowances allocated turned out to be 4 % more than actual emissions, referred to as allowance 'over-allocation' (Ellerman and Buchner 2008).

After trial Phase I, the real test came in Phase II (2008–2012). Phase II started with an ample issue of allowances for the year 2008 (without a buffer of allowances from Phase I, since banking allowances from Phase I to Phase II was not allowed). Phase II aimed at an emissions cap for 2012, set at 6.5 % below 2005 CO_2 emission levels. The allowance price hit a peak of € 35 in July 2008. This price did not so much reflect the immediate necessity to abate CO_2 emissions and its marginal cost, but rather the expectations about the scarcity of allowances in the years to come. Those expectations were based (a) on the expected rate of output growth in the carbon-intensive industry brought under the regime of the EU ETS, and (b) on the prevailing ideas of how high marginal CO_2 emission control cost would run up by the time that the emissions cap would incentivize firms to take emission reduction measures. European industry officials had repeatedly expressed their concern that the CO_2 price would become too high (Grubb and Neuhoff 2006) and this seemed to become reality. Policymakers and traders saw a trajectory with expected allowance prices of € 25–€ 35 in 2010, and of € 35–€ 50 in 2020, as a plausible scenario (Point Carbon 2008: 31).

However, the fall of Lehman Brothers in September 2008 initiated a banking crisis that lasted until March 2009. It brought an economic bust that triggered an implosion of sales and output of carbon-intensive industries under the EU ETS. In 2009, the CO_2 emissions of those industries were 11.6 % lower than in 2008, which was even considerably below the emissions cap for 2012. This caused a sudden growth in so-called 'surplus' allowances that were not used to offset CO_2 emissions in 2009. This surplus of unused allowances grew further in the years thereafter, since production only recuperated hesitatingly (with industrial production of the EU in 2015, which can be used as an indicator of the output of carbon-intensive industries, being equal to its level in 2005). As a consequence, the allowance price sank to € 9 early 2009.

After the recovery from the banking crisis in the US started to improve the global macro-economic situation, the EU fell into a debt crises in 2010 which brought even more years of economic difficulties. The output of the industry covered by the EU ETS recovered only slowly and the buffer of unused allowances increased even further, up to 2 billion metric tonnes of CO_2-equivalent at the end of

2012, which is more than needed to offset the emissions of a full year. In the first year of Phase III (2013–2020), aiming at an emissions cap for 2020 set at 21 % below 2005 CO_2 emission levels, the allowance price sank to € 3. At the end of 2013 the allowance surplus had increased to 2.1 billion tonnes. In 2014 and 2015, the allowance surplus has not changed much and the carbon price stood around € 5 early 2016.

The events of the past seven years have drastically changed the expectations regarding the EU ETS. Concerns by the industry about a too high carbon price were overshadowed by worries among European legislators that CO_2 emissions are priced too low. Such disappointments easily feed slumbering feelings of resentment against market-based policy instruments. How much more does it take to awaken them and call for ending the 'failed' experiment of emissions trading in Europe? Some statements at the beginning of this section reflect this sentiment. But are they based on economic logic?

The EU ETS was obviously adopted in the belief that the reduction of CO_2 emissions through an allowance market would work so that the emissions cap for 2020 would be achieved. Can the succession of events up to early 2016, in particular the lower than foreseen allowance price, be interpreted as a threat to realizing that objective? Certainly not; the contrary is true. The past years of unexpectedly low CO_2 emissions, due to output levels that were below initial expectations, have mightily increased the allowance buffer. That surplus of allowances has actually made it easier to realize the target for cumulative CO_2 emissions, up to and including 2020, at surprisingly low costs. One can predict that investments in emission reduction, that in a scenario of steady growth in output would have occurred in Phase III, will not be undertaken because it will turn out that there is no need for them. They can be viewed as a block of potential investments in greenhouse gas emission reduction available for the years after 2020. That insight can be helpful to mobilize the political support for a high ambition level in setting the emission targets for 2030, the last year of Phase IV, and beyond.

The seven year experience of a lower than expected carbon price has inspired a long list of publications either proposing amendments to the EU ETS or proposing alternative instruments. Not only several economists and lawyers have done so (e.g. Mulder 2015; Avi-Yonah and Uhlmann 2009), but in particular *The Economist* has proven to be a leading popular medium in this literature. This weekly magazine has never made a secret of its preference for a carbon tax over carbon emissions trading (see e.g. *The Economist* 2013). The price history of the EU ETS has only affirmed its view on the matter. Indeed, it seems so straightforwardly simple. A tough tax of, say, € 20 per tonne of CO_2 emission set in 2005 and kept constant, would have given the industry now covered under the EU ETS the incentive to abate CO_2 emissions up to the level where its marginal abatement cost equals the tax per tonne of emitted CO_2. A carbon tax would over the past decade have delivered emission reductions far above the little that has now been realized by the EU ETS. Cumulated CO_2 emissions would have been spectacularly lower.

It is not to be denied that the European Commission in planning CO_2 emission reduction targets has fallen victim of its imperfect foresight of the future and did

grossly overestimate economic growth. But it is also true that *The Economist* and academics focusing on the unforeseen recession tell only half of the story. People who know to have imperfect foresight should also consider the other half of the story, which is an unforeseen boom. If this would have occurred, (a) the demand for allowances under the EU ETS would have been higher than the official prognosis, (b) the allowance price would have climbed to a level higher than the carbon tax, and (c) abatement of CO_2 emissions would also have been higher, while CO_2 emissions would have been on target in 2020 thanks to the emissions cap. Under a carbon tax, however, accumulated emissions of the industry in 2020 would be above the target due to a too low tax rate in the face of the unexpected strong increase in output and associated unabated emissions in Phases II and III.

So when the full story is told about these two economic instruments operating in a world of imperfect foresight, the overall conclusion must be that a carbon tax delivers either a better or worse result than the target level for accumulated emissions, while an emissions trading scheme based on cap-and-trade can only end with a better result. With imperfect foresight, cap-and-trade is thus a more effective instrument to mitigate climate change than a carbon tax.

Should the carbon tax then perhaps be preferred because of its economically less painful side effects on such variables as industry output, employment and profits? In making the comparison we shall place a cap-and-trade design of the EU ETS against a tax per tonne of CO_2 emission. In the early design of EU ETS, as originally proposed by the European Commission, all allowances would be handed out for free. In an economically naive view of the world that may look as if using such free allowances to offset emissions is without a cost. However, such a belief is a mistake. The seemingly free allowances actually have an opportunity cost. There are two ways, based on economic theory, to explain this cost. The first explanation is that using the allowance for producing a unit of output will not be available for producing a next product, for which the associated allowance would have to be bought. The second explanation is that the allowance that is used to enable production cannot be sold for money. In both explanations using the allowance for production is an economic sacrifice and therefore a cost of production. In this respect there is an analogy with a tax on CO_2 emissions which is, in the end, also part of the cost of producing output.

However, next to the similarity in being a cost of output, there is also a difference. In a cap-and-trade scheme the allowances to offset emissions are, on the one hand, a cost of output that therefore results in a higher price of output. But on the other hand, if the allowances serving as an input to output have been handed out for free, they are owned by the firm and as supplying the input yields a revenue equal to the opportunity cost it therefore adds to the total profit from output. By contrast, a tax on CO_2 emissions is also a cost of output, but the tax yield goes to the public coffer. Hence it is not difficult to understand why industry preferred the original design of the EU ETS as a cap-and-trade scheme with free permits over a CO_2 emission tax. The aforementioned difference is of utmost importance in times of economic depression when the price of output falls. The marginal firms in an industry may survive under cap-and-trade with free allowances as long as the loss

due to the lower price of output does not exceed the yield of supplying the allowances, whereas a marginal firm under a carbon tax regime would succumb because of the net loss inflicted by the low price of output. The economic interests of the industry, measured in terms of output, profit and employment, is better served by cap-and-trade with free allowances than it is by a carbon tax.

In addition, there is a macro-economic interest in having a cap-and-trade scheme for restricting carbon emissions rather than a carbon tax. In times of recession there will be a low allowance price that does not weigh heavy on the shoulders of energy-intensive industries. In such an economic bust, output shrinks and the use of fossil fuels and carbon emissions go down. The allowance demand function shifts downward and the market price of allowances falls. The necessity to abate emissions as well as the need to obtain allowances for offsetting emissions decrease. As a consequence, the cost of complying with climate policy gets lower during the bust. For sectors exposed to international competition, the so-called 'exposed sectors', it improves international competitiveness, for example against countries without a policy of carbon emission reduction, so that the scheme works out as an economic stimulus. In the past years of economic stagnation, the enormous drop in CO_2 prices has practically annihilated the compliance cost for many firms under the EU ETS. In an economic boom, however, the system works in the opposite direction, mitigating the upswing. When the economy starts to grow again, the allowance price will increase and the industry will be better able to bear it. The consequence is that emissions trading works 'countercyclical': allowance prices increase when economic times are good and decrease when times are bad. A cap-and-trade scheme thus functions as an economic stabilizer, a property welcomed by macro-economists. In a carbon tax this desirable feature is lacking.

Advocates of a carbon tax usually envisage a constant or increasing tax rate. How likely would such a scenario have been over the past decade from a political-economy point of view? Around 2000, the industry lobbied hard against the creation of an EU ETS, arguing that a stringent cap on total emissions would cause unduly high cost of production for the industry and undermine its international competitiveness, while it would also lead to a loss of market share, shrinking employment and carbon leakage. In the battle of wits that last point in particular was a good 'hit'. Carbon leakage is the technical term for the phenomenon that carbon emitting industries would lose customers or might even see firms migrating to non-EU countries with weak carbon reduction policies (e.g. countries without a carbon pricing regime). It could even be imagined that CO_2 emission reduction policy in Europe would result in a worldwide increase of CO_2 emissions instead of a decline. The lobby was highly successful in being heard, leading to various concessions in the design of the EU ETS that was adopted in 2002 (such as free allowances, among others for exposed sectors and newcomers). Not really a political environment where a tax proposal, with its economic disadvantages for the energy-intensive sectors compared to cap-and-trade with free allowances, would have made much chance of coming through. In fact, the idea to introduce carbon emissions trading was born out of the political necessity to come up with an alternative instrument after the Council of Ministers in 1993 had rejected the proposal for a European carbon tax as being economically too painful. It is unlikely, if

not absurd, to think that less than ten years after 1993 a European carbon tax would have become politically feasible. There is a fair chance that the outcome would have been a political stalemate and further years of non-policy. That is a world far away from the scenario of a constant or increasing carbon tax over the past economically contagious years.

Nevertheless, suppose that not the EU ETS Directive, but a carbon tax would have been adopted in 2002. If EU Member States would have agreed on a carbon tax at a level high enough to prevent growth of CO_2 emissions, given the expectation of production growth, there would have been serious pressure from the industry after September 2008 to lower the tax rate in the face of shrinking demand for output and the threat of carbon leakage due to loss of international competitiveness. Taking into account the probability that a carbon tax will be adjusted to the political-economic demands of the time, (namely a lower tax in an economic bust and a higher tax in an economic boom) makes that a carbon tax in its time path will show more resemblance with that of the carbon price in an emissions trading scheme. In a similar vein, the foreseen introduction in the EU ETS of a so-called 'market stability reserve' (MSR) in 2019—that will serve to reduce the allowance auction volume in case the allowance surplus is considered 'too big' and to release extra allowances if the surplus is 'too small'—is a step that makes the EU ETS look somewhat more similar to a carbon tax in its economic impact.

The conclusion is that in blueprint cap-and-trade is more effective in reducing emissions than a carbon tax, whereas in the real world the difference in time paths of carbon price, annual emission abatement and accumulation of CO_2 emissions would be considerably less between the two instruments than calculations show in models that abstract away from the political context of instrument choice.

3 Second Misconception: 'The Objective of the EU ETS Is to Stimulate Low-Carbon Technologies'

Some authors and actors describe the EU ETS in such a way that it seems as if the objective of the EU ETS is to stimulate low-carbon technologies. For instance

> The EU ETS is the main instrument of European climate policy, and many policymakers envisage it as a driving force of the EU's transition to a low-carbon economy. By putting a price on emissions, the scheme is expected to encourage heavy polluters to develop new low-carbon technologies. (Calel and Dechezleprêtre 2013: 1).

> The European Union commenced the pilot phase of the European Union Emissions Trading System (EU ETS) in 2005 with the intent to enhance the adoption of existing low-carbon technologies and the development of new ones by putting a price on CO2 emissions. (Anderson et al. 2010: 1).

Such descriptions of the objective of the EU ETS are certainly not wrong but they are incomplete, both from a legal as well as from an economic point of view.

The primary legal objective of the EU ETS is '(…) to promote reductions of greenhouse gas emissions in a cost-effective and economically efficient manner' (Article 1, EU ETS Directive 2003/87/EC). Some legal scholars argue that reducing greenhouse gases is primary, while doing so cost-effectively is secondary; others argue that the entire objective stated above is primary (Squintani et al. 2012). To this end, Article 3 of the EU ETS Directive defines an emission right as '(…) an allowance to emit one tonne of carbon dioxide equivalent during a specified period, which (…) shall be transferable (…)'. This is in line with environmental economic theory which conceptualizes emissions trading as an instrument to reduce emissions at the lowest possible cost (e.g. Dales 1968).

The secondary or implied objectives are (a) low-carbon technological innovation (e.g. Recital 20 in Preamble of EU ETS Directive to 'encourage the use of more energy-efficient technologies' and Article 10 of the EU ETS Directive on using part of new entrants reserve to stimulate innovative low-carbon technologies) as well as (b) preventing carbon leakage and protecting the competitiveness of internationally operating industries (regulated e.g. in Article 10 of the EU ETS Directive that provides for free allowances for companies that compete with firms outside EU).

Legally qualifying low-carbon innovation as a secondary or implied objective also makes economic sense. Emitters that surpass their emissions cap have a menu of options at their disposal. They can buy allowances to cover the emissions deficit (or pay the fine for non-compliance) or reduce emissions within the company itself, for instance using new technologies that save energy. The latter option underlines that low-carbon technologies are indirectly stimulated by an emissions trading scheme that primarily aims to reduce emissions at the lowest possible cost. Such technologies are in fact necessary to be able to meet the ambition of the EU to reduce greenhouse gas emissions by 80–95 % below 1990 levels in 2050 (European Commission 2011). This also explains the legitimate focus of researchers, such as the ones cited above, to study the implied effect of the EU ETS on low-carbon innovation.

Clarity on the primary aim of the EU ETS is of utmost importance for the political discussion on the level of the allowance price. Many observers criticize the allowance price because its current level of around € 5 is much lower than initially predicted. About a decade ago, as indicated before, policymakers and traders expected allowance prices of € 25–€ 35 in 2010 and of € 35–€ 50 in 2020 (Carbon 2008: 31).

Such a low allowance price is in conformity with the primary objective of cost-effectiveness and the secondary objectives of industry protection and carbon leakage prevention. Moreover, as was argued in the previous section, a low allowance price is an advantage in times of economic recession: it does not weigh heavy on the shoulders of big and small energy consumers. When the economy starts to grow again, the allowance price will increase and consumers will be better able to bear it.

However, a low allowance price is in contradiction with the objective of low-carbon innovation, which will not be sufficiently stimulated by such a low price. The latter view, focusing on the secondary or implied objective of the

EU ETS, has led the EU to adopt new measures that indirectly stimulate the allowance price.

First, the emissions cap for ETS sectors will be lowered by 2.2 % each year from 2021 onwards. This is indeed the right thing to do, not only because emissions trading is a quantity-based instrument (and not a price-based instrument like a carbon tax), but also because a declining emissions cap is necessary to reach the deep emissions cuts foreseen by the EU in 2050 in light of the expected climate damage. The EU has resisted calls for directly intervening in the market price, which would artificially raise the allowance price under the agreed emissions cap and thus lead to inefficiencies in the scheme.

Second, a so-called 'market stability reserve' (MSR) will be established that reduces the allowance auction volume in case of an allowance surplus (European Council 2014). If the allowance surplus is 'too big' in a certain year (more than 833 million allowances), 12 % of this surplus will be placed in the reserve by reducing the auction volume with a corresponding amount of allowances in the year thereafter. If the allowance surplus is 'too small' (less than 400 million allowances), 100 million allowances will be released from the reserve. Moreover, the auctioning of 900 million allowances will be postponed for a number of years, referred to as 'back-loading' (European Commission 2014). These 900 million 'back-loaded' allowances will be put into the MSR, which will become operational in 2019 (European Parliament 2015). Based on the previous section, we argue that the MSR and 'back-loading' are of no use as instruments to achieve the target for total accumulated emissions. It remains to be seen whether it will be effective as an instrument to lift up the carbon price in the near future and, as a consequence, to stimulate low-carbon technology. To raise the carbon price, a permanent deletion of a number of allowances would be more effective (and probably also less complex) than putting them temporarily in a reserve.

In sum: the primary objective of the EU ETS is not to stimulate low-carbon technologies, but to reduce greenhouse gas emissions at the lowest possible cost. Nevertheless, low-carbon innovation is an important secondary objective or implied effect, because companies that face a declining emissions cap might also find it financially attractive to reduce emissions within the firm itself, for instance by developing and/or applying novel abatement technologies.

That said, the effectiveness of emissions trading as a stimulus for the development of low-carbon technology lacks a solid base in the economic literature. In the analytical models regarding technical progress, improvement in emission control technology is defined as achieving a downward shift in the total and marginal cost of the emission abatement function. The generally accepted conclusion, derived from simple models in which output is held constant, is that a carbon tax is a stronger incentive for research and development of cleaner and lower cost control technology than direct regulation based on emission performance standards. In such a scheme of direct regulation, assuming a given level of output, the performance standard determines the required level of emission abatement. The installation and operation of innovative control technology brings down the cost of realizing the required amount of emission control. Under a carbon tax there is, next to this saving

on abatement cost, in addition a saving on the amount of taxes paid for carbon emissions, because the innovating firm increases its abatement level now that the innovative control technology has brought down the cost of doing so. Therefore, the net benefits from installing innovative low-carbon technology are higher under a constant carbon tax than they are under a performance standard, given the assumption of an unchanged level of output.

At first glance, one might think that emissions trading schemes will deliver a financial stimulus to low-carbon technology comparable to a carbon tax, but it is not like that. The installation of innovative emission control technology that has lower costs leads to a lower demand for emission allowances and consequently to a decrease in the market price of allowances. It implies that firms that did not install the new technology benefit from buying the allowances that they need at a lower price. Now that allowances have become cheaper, the non-adopter will buy more allowances and cut back on relatively expensive emission control. The theoretical consequence is that the advantage for firms who adopt the innovative low-carbon technology compared to non-adopters—the 'innovation incentive'—is even less under cap-and-trade than it is under direct regulation (Keohane 1999; Requate and Unold 2003), although the scarce empirical literature on this issue is still inconclusive (e.g. Popp 2003; Taylor et al. 2005; Schmidt et al. 2012).

One should note, however, that the above ranking has been derived in a static model, based on the assumption of constant output per firm. In a world where economic growth can be faster than foreseen when the targets for accumulated carbon emissions were planned, it can occur that in a cap-and-trade scheme the carbon price runs up to a level (far) above the constant carbon tax and that the upward drive in price even surpasses the carbon price depressing effect of ongoing innovations in emission control technology. In such a dynamic economy, cap-and-trade will outperform the constant carbon tax, both in adequately restricting carbon emissions and in stimulating low-carbon technology. In an economy more sluggish than forecasted, the constant carbon tax is the superior instrument with respect to both criteria.

We conclude that there is quite a gap between the official secondary objective of low-carbon innovation in the EU ETS and the implications derived from analytical models, which classify cap-and-trade as inferior to direct regulation and rank a carbon tax that is held constant over time as the best instrument to stimulate low-carbon technologies. However, the recent switch of the EU ETS to a kind of 'benchmark-and-trade' system, to be discussed in the next section, gives the European carbon market more features of performance standard rate trading than of cap-and-trade. The static analytical models predict that this modification will strengthen the economic incentive that drives progress in CO_2 control technology. Remember, however, that this result only holds under the assumption of a world with constant output. Moreover, even if this assumption would hold in reality, the advantage of stimulating innovation comes at the expense of cost inefficiency in the current design of the EU ETS, as we will explain below.

4 Third Misconception: 'The EU ETS Is a Cap-and-Trade Scheme'

Various authors and actors describe the EU ETS as a cap-and-trade scheme. For instance

> (…) Europe's flagship environmental programme, the Emissions Trading System (ETS), (…) is a cap-and-trade scheme in which permits to emit carbon (…) are allocated to firms and can then be traded between them. (The Economist 2013: 1).

> (…) the EU has adopted a flawed cap-and-trade system. (Avi-Yonah and Uhlmann 2009: 50).

> The European Union Emissions Trading System (EU ETS) is a progressive cap-and-trade system aimed at reducing global greenhouse gas emissions emitted from Europe. (Peters 2015: 1)

Again these descriptions are not wrong but they are imprecise, from a law and economics point of view. It is too simple to label the EU ETS as a cap-and-trade scheme. The EU ETS is, in fact, a hybrid of two distinct environmental policy instruments, namely permit trading (also referred to as allowance trading or cap-and-trade) and credit trading (also referred to as performance standard rate trading or output-based allocation.

We have argued this for a number of years (e.g. Nentjes and Woerdman 2012), and continue to do research on this complex regulatory mixture (Woerdman and Nentjes 2016), but the fact that the EU ETS is a hybrid is still not well understood in the literature on emissions trading in Europe. This design amalgam is important to understand, since it sheds light on the inefficiencies in the rules of the trading scheme, including those rules that have been amended recently.

Permit trading, also referred to as allowance trading or cap-and-trade, is a conceptual variant of emissions trading that refers to the trading of emission entitlements under an emissions cap. A permit trading system imposes a cap on the annual emissions of a group of companies for a certain period of time. Emission rights, called 'allowances', are allocated to established companies for this period. The allowances are allocated either for free or through an annual sale by auction. The allowances are tradable. Newcomers and companies seeking to expand production must purchase allowances from established companies or from a government reserve. A company closing down a plant can sell its allowances.

Credit trading, also referred to as performance standard rate trading (PSR) or output-based allocation (OBA), is a different conceptual variant of emissions trading that refers to the trading of emission entitlements based on a performance standard. Such a system of tradable reduction credits is based on a government-mandated emissions standard adopted for a group of companies. The emissions standard dictates permitted emissions per unit of energy consumption or per unit of production output. In this system, emission reduction credits can be earned by emitting less than prescribed by the emissions standard. These credits can then be sold to companies who can use them to compensate their emissions in excess of the emissions standard applying to them. If the economy grows, the

supply of credits also increases because companies do not operate under an absolute emission ceiling but have to observe the relative emissions standard. An energy-intensive company that expands production, or a newcomer entering the industry, therefore has a right to new emissions, as long as he obeys the emissions standard. This means that emissions will grow in absolute terms. A company closing down loses its credits.

There is a lack of full consensus in the emissions trading literature about which design variant is 'superior', for instance because studies apply different performance criteria, models and assumptions. However, most authors agree that permit trading outperforms credit trading in terms of effectiveness and efficiency (for an overview of this literature see, e.g. Nentjes and Woerdman 2012; Tietenberg et al. 1999).

Permit trading is more effective than credit trading in reaching the absolute emission targets of companies. Permit trading imposes an absolute limit on total emissions, which also should become more stringent over the years. Under credit trading, however, when industrial production increases the emissions of companies rise as well. The only way to deal with this problem is to strengthen the emissions standard (ad hoc or perhaps automatically), but energy-intensive industries are likely to lobby against such a stricter emissions requirement.

Permit trading is also more efficient than credit trading. In an emissions trading system based on emissions caps, each unit of emissions has a price. In case of allowance auctioning, the emitter has to buy allowances and obviously pays for each unit of emissions. In case of free allocation, allowances have opportunity costs, equal to the allowance price. Using the allowances to cover the emissions, the emitter foregoes the opportunity to sell the allowances and thus misses sales revenues. Since each unit of emissions has a price in a cap-and-trade scheme, there is an incentive to examine all emission reduction possibilities and apply the least-cost option.

In a credit trading system, there is a credit price. The received amount of money for credits sold is equal to the sum paid by companies that exceed the emissions standards to purchase the credits. However, the emissions within the limits set by the emissions standard remain without a price. For the group of companies as a whole, the cost of the permitted emissions is nil. Free credits do not have opportunity costs: if a company stops emitting he has no credits to sell, because he will loose its credits. Credits can only be earned through reducing emissions per unit of energy or output. Economising on fuel input or slowing production does not earn any credits. Total emission reduction costs are therefore higher compared to permit trading.

To be able to judge the efficiency of the EU ETS, it is paramount to see that the complex rules of this scheme contain elements of both permit trading and credit trading.

The first mix of permits and credits can be seen in the 'benchmarking' rules to allocate free allowances, making the EU ETS a kind of 'benchmark-and-trade' system. Since 2013, the default method for allocating allowances is through auctioning them (e.g. to power companies), but various energy-intensive industries still

receive allowances free of charge to prevent carbon leakage: the undesirable moving of industrial activity, and thus emissions, to non-EU countries with weak carbon reduction policies (e.g. countries without a carbon pricing regime). Initially, free allowances were allocated based on historical emissions, but from 2013 onwards, free allocation of allowances will take place on the basis of a carbon standard per unit of production multiplied by production in 2005 (or the average for 2005–2007 if this is higher). Such an emissions standard, referred to as 'ex-ante benchmark' in Article 10a of the amending EU ETS Directive, is determined based on the average emissions of the 10 % installations with the lowest carbon emissions per unit of product or energy output in an industrial sector in the years 2007–2008. These complex rules not only come down to handing out free allowances based on low-carbon performance, but also changed the initial allocation of free allowances in the advantage of companies that had the strongest growth in output in the decade up to 2005. One might argue that a prerequisite for credit trading is brought into the EU ETS using an emissions standard ('benchmark') to calculate the allocation of free allowances, but 'benchmarking' does keep in place the cap on emissions per company, so that the cap-and-trade features of EU ETS are basically preserved here. This picture changes, however, when one considers the next policy element.

The second mix of permits and credits can be seen in the altered rules for newcomers and expansions, which have a serious distortive economic impact. In a pure cap-and-trade scheme, newcomers and companies that want to expand need to buy allowances from established companies or from a government reserve. In the EU ETS, however, newcomers as well as industries expanding their production capacity (in excess of 10 %) will be allocated allowances for free. A newcomers' reserve has been created to facilitate this, equalling 5 % of the total number of allowances. In addition, allowances in a pure cap-and-trade scheme do not expire when an installation closes down or when its capacity is reduced. In the EU ETS, however, allowances need to be surrendered in case of plant closure or significant decline in production capacity. The consequence of these rules for newcomers, expansions and closures, which resemble those of a credit trading system, is that they lead to the following two inefficiencies (Woerdman 2015).

First, if companies would actually keep their allowances in case of installation closure, it would be more attractive to shut down old, climate-unfriendly plants since the allowances could then be sold on the carbon market. The rule in the EU ETS that companies have to surrender free allowances in case of closure is inefficient since it undermines this desirable incentive. It makes the closure of dated, usually carbon-intensive plants less attractive (Ellerman 2007).

Second, when product prices are so low that (variable) production costs cannot be covered anymore, a company would normally shut down its installations and leave the market. As a result of the expansion and closure rules of the EU ETS, however, companies that make losses maintain their production capacity in order to continue receiving free allowances which can be sold on the carbon market. In the (unlikely) case of a very high allowance price, it would even be profitable to invest in production capacity only to obtain allowances to sell. Investing in or maintaining capacity which is not deployed for production purposes is inefficient and constitutes

a social waste (Nentjes and Woerdman 2012). To counter this effect the EU has adopted additional rules which stipulate that if the level of production is cut back below a certain percentage of production capacity, the number of free allowances will be adjusted downwards. For a novel and more detailed analysis of all rules for newcomers and expansions in the EU ETS, the reader is referred to Woerdman and Nentjes (2016) where it is concluded that these rules contain various perverse incentives.

The third mix of permits and credits can be seen in the recently proposed rules for basing the free allocation of allowances on companies' changing production levels. So far, the European Commission has resisted the industry lobby for 'intended' emissions caps for companies, but the implementation problems of the EU ETS and the corresponding political desire to reform the scheme have already opened the political window for sub-optimal ex-post corrections of the allocated allowances. The European Council stated that: 'Future allocations will ensure better alignment with changing production levels in different sectors' (European Council 2014: 2). Although the detailed rules still need to be elaborated, this is likely to imply that companies receive more allowances when their production and thus emission levels increase, which does not reflect the social costs of the additional emissions. In addition, there will probably be less abatement options for companies, raising total emission reduction costs vis-à-vis cap-and-trade, since slowing production will not earn any allowances. This recent policy development shows that the EU ETS is increasingly becoming a credit trading scheme for the exposed industries, although the gradual expansion of auctioning allowances partly mitigates this.

In sum: the EU ETS is not a pure cap-and-trade scheme, but a complex hybrid with elements of permit trading (cap-and-trade) and credit trading (performance standard rate trading). This leads to various inefficiencies, including the situation that companies could maintain unprofitable production capacity in order to continue receiving free allowances. The rules as they are also make the closure of old plants less attractive compared to a pure cap-and-trade scheme.

5 Conclusion

The EU ETS faces multiple implementation problems, but the debate is blurred, over-simplified and sometimes even misguided. According to recent empirical research, '(…) there remains an insufficient level of understanding about the ETS in some of the EU-28 member states. Not all stakeholders are fully informed about or fully understand exactly how the ETS works.' (Fujiwara et al. 2015: 17). In a different way, our chapter confirms these findings. There appears to be confusion about what the EU ETS should deliver, whether it is effective in doing so, and how the scheme should be characterized. In our chapter we have corrected a number of these misconceptions about the EU ETS, which should contribute to a more nuanced and more precise debate on emissions trading in Europe.

The first misconception that we confront is the rather popular idea that a carbon tax would be more reliable and effective than the EU ETS to bring down CO_2 emissions in Europe. It is actually the other way around. In a long-lasting economic 'boom' a carbon tax may result in cumulated CO_2 emissions above the target level of the EU, which will not happen in a trading scheme based on emission caps. Moreover, in years of economic 'bust' a carbon tax has negative macro-economic impacts, which may result in carbon leakage undoing the success in keeping cumulated CO_2 emissions in Europe below the target. A cap-and-trade scheme, however, functions as an economic stabilizer: allowance prices increase when economic times are good and decrease when times are bad. Therefore, the EU ETS is the most effective and efficient instrument to cut back CO_2 emission levels to the targets set for 2020 and beyond.

The second misconception that we correct is the emerging view that the principal objective of the EU ETS is to stimulate low-carbon technologies. The objective of the EU ETS is, first and foremost, to reduce greenhouse gas emissions at the lowest possible cost. Low-carbon innovation is an important secondary objective or implied effect, because companies that face a declining emissions cap might also find it financially attractive to reduce emissions within the firm itself, for instance by developing and applying novel abatement technologies. That said, we have also argued that a carbon tax is probably more effective in stimulating research, development and adoption of low-carbon technology than the EU ETS, because innovation drives down the allowance price which subsequently weakens the innovation incentive. Nevertheless, in a buoyant output market it might be the other way around.

The third misconception that we put right is the widespread notion that the EU ETS is a cap-and-trade scheme. The EU ETS is, in fact, a hybrid scheme that combines elements of permit trading (cap-and-trade) and credit trading (performance standard rate trading). This leads to inefficiency in the way that CO_2 emissions are restricted. Instead of lowering the share of carbon-intensive consumption in total income at relatively low cost, as would occur under cap-and-trade, the emissions of carbon-intensive production sectors under the EU ETS are abated at relatively high cost, for instance by installing and operating additional emission control technology and by postponing the closure of old carbon-intensive plants.

References

Anderson, B., F. Convery, and C. Di Maria. 2010. *Technological Change and the EU ETS: The Case of Ireland*. Dublin: University College Dublin.

Avi-Yonah, R.S., and D.M. Uhlmann. 2009. Combating Global Climate Change: Why a Carbon Tax is a Better Response to Global Warming than Cap and Trade. *Stanford Environmental Law Journal* 28(3): 3–50.

Backhaus, J.G. 2005. Towards an Ideal Economic Analysis of a Legal Problem. In *Elgar Companion to Law and Economics*, ed. J.G. Backhaus, 465–472. Cheltenham: Edward Elgar.

Calabresi, G., and A.D. Melamed. 1972. Property Rules, Liability Rules, and Inalienability: One View of the Cathedral. *Harvard Law Review* 85(6): 1089–1128.

Calel, R., and Dechezleprêtre, A. 2013. *Has the EU ETS Induced Low-carbon Innovation?*, Green Growth Knowledge Platform, 22 July 2013.

Coase, R.H. 1960. The Problem of Social Cost. *Journal of Law and Economics* 3: 1–44.

COM. 2015. *337, Proposal for a Directive of the European Parliament and of the Council Amending Directive 2003/87/EC to Enhance Cost-effective Emission Reductions and Low-carbon Investments, 15-7-2015*. Brussels: European Commission.

Dales, J.H. 1968. *Pollution, Property and Prices: An Essay in Policy-Making and Economics*. Toronto: Toronto University Press.

Directive 2009/29/EC of the European Parliament and of the Council of 23 April 2009 Amending Directive 2003/87/EC so as to Improve and Extend the Greenhouse Gas Emission Allowance Trading Scheme of the Community. *OJ* 2009 L.140/63–87.

Directive 2003/87/EC of the European Parliament and the Council of 13 October 2003 Establishing a Scheme for Greenhouse Gas Emission Allowance Trading within the Community and Amending Council Directive 96/61/EC. *OJ* 2009 L.275/32–46.

Economist. 2013. 'ETS, RIP?', *The Economist*, 20 Apr 2013.

Ellerman, D. 2007. New Entrant and Closure Provisions: How Do They Distort? *The Energy Journal* 28: 63–78.

Ellerman, D., and B. Buchner. 2008. Over-allocation or Abatement? A Preliminary Analysis of the EU ETS Based on the 2005-06 Emissions Data. *Environmental & Resource Economics* 41(2): 267–287.

European Commission. 2011. *A Roadmap for Moving to a Competitive Low Carbon Economy in 2050*, COM(2011) 112 final, 8.3.2011, Brussels: European Commission.

European Commission. 2014. *Commission Regulation (EU) No 176/2014 of 25 February 2014 Amending Regulation (EU) No 1031/2010 in Particular to Determine the Volumes of Greenhouse Gas Emission Allowances to be Auctioned in 2013–20*. *OJ* 2014 L.56/11–13.

European Council. 2014. *Conclusions on 2030 Climate and Energy Policy Framework*, EUCO 169/14, 24.10.2014, Brussels: European Council.

European Parliament. 2015. *ETS Market Stability Reserve: MEPs Strike Deal with Council*, European Parliament Press Release 06–05-2015.

Faure, M., and M. Peeters (eds.). 2008. *Climate Change and European Emissions Trading: Lessons for Theory and Practice*. Cheltenham-Northampton: Edward Elgar.

Fujiwara, N., et al. 2015. *How Do Stakeholders View the EU ETS? Diversity and Differentiation of Interests*, PolImp Working Document No. 2, May 2015.

Greenpeace. 2013. *New Report Finds Failing EU Carbon Market Threatens Effectiveness of 2030 Climate Proposal*s, Press release 11 June 11 2013, Brussels: Greenpeace, http://www.greenpeace.org/eu-unit/en/News/2013/Ecofys-Briefing/.

Grubb, J., and K. Neuhoff. 2006. Allocation and Competitiveness in the EU Emission Trading Scheme: Policy Overview. *Climate Policy* 6(1): 7–30.

Keohane, N.O. 1999. *Policy Instruments and the Diffusion of Pollution Abatement Technology*. Mimeo: Harvard University.

Lang, C. 2013. *The EU Emissions Trading Scheme has failed: "Time to scrap the ETS"*, REDD-Monitor, 16 Apr 2013. http://www.redd-monitor.org/2013/04/16/the-eu-emissions-trading-scheme-has-failed-time-to-scrap-the-ets/.

Mulder, A.J. 2015. *CO_2 Emissions Trading in the EU: Models and Policy Applications*, Dissertation, University of Groningen.

Nentjes, A., and E. Woerdman. 2012. Tradable Permits versus Tradable Credits: A Survey and Analysis. *International Review of Environmental and Resource Economics* 6(1): 1–78.

Peters, V.H. 2015. A Legal Obligation to Mitigate Greenhouse Gas Emissions from Agriculture: A Challenge to the European Union's Emissions Trading System and the EU Member States with the Largest Agricultural Impact, *19 UCLA J. Int'l L. Foreign Aff. 213.*

Point Carbon. 2008. *Carbon 2008: Post-2012 is Now*. Oslo: Point Carbon.

Popp, D. 2003. Pollution Control Innovations and the Clean Air Act of 1990. *Journal of Policy Analysis and Management* 22(4): 641–660.

Requate, T., and W. Unold. 2003. On the Incentives to Adopt Advanced Abatement Technology: Will the True Ranking Please Stand Up? *European Economic Review* 47: 125–146.

Schmidt, T.S., M. Schneider, K. Rogge, M.J.A. Schuetz, and V.H. Hoffmann. 2012. The Effects of Climate Policy on the Rate and Direction of Innovation: A Survey of the EU ETS and the Electricity Sector. *Environmental Innovation and Societal Transitions* 2: 23–48.

Squintani, L., J.M. Holwerda, and K.J. de Graaf. 2012. Regulating Greenhouse Gas Emissions from EU ETS Installations: What Room is Left for the Member States? In *Climate Law in EU Member States, towards National Legislation for Climate Protection*, ed. M. Peeters, and M. Stallworthy, 67–88. Cheltenham UK: Edward Elgar Publishing.

Taylor, M.A., E.S. Rubin, and D.A. Hounshell. 2005. Regulation as the Mother of Innovation: The Case of SO_2 Control. *Law and Policy* 27(2): 348–378.

Tietenberg, T., Grubb, M., Michaelowa, A., Swift, B., Zhang, Z.X. 1999. *International Rules for Greenhouse Gas Emissions Trading*, UNCTAD/GDS/GFSB/Misc.6, Geneva: United Nations Conference on Trade and Development (UNCTAD).

Woerdman, E. 2015. The EU Greenhouse Gas Emissions Trading Scheme. In *Essential EU Climate Law*, ed. E. Woerdman, M.M. Roggenkamp, and M. Holwerda, 43–75. Cheltenham: Edward Elgar.

Woerdman, E., and Nentjes, A. 2016. '*Emissions Trading Hybrids: The Case of the EU ETS', Working Paper*. University of Groningen: Department of Law and Economics (revised & resubmitted to the *Review of Law and Economics*).

Erratum to: The Law, The Economy, The Polity Jürgen Backhaus, A Thinker Outside the Box

Jean-Michel Josselin, Alain Marciano and Giovanni Battista Ramello

Erratum to: Chapter "The Law, The Economy, The Polity Jürgen Backhaus, A Thinker Outside the Box" in: A. Marciano and G.B. Ramello (eds.), *Law and Economics in Europe and the U.S.*, The European Heritage in Economics and the Social Sciences, DOI 10.1007/978-3-319-47471-7_1

The book was inadvertently published with incorrect information in Chapter 1, which has been corrected now as below:

In Chap. 1, P. 4, last line in first paragraph of Section 2: the sentence is changed to "... restrict its attention to behaviors that can be...", in line 6 in second paragraph: the close quote position in the text "pure theory of government (1954)" has been changed and it should read as "pure theory of government" (1954).

In P. 6, line 6: "of" is deleted before the word "rats".

The updated original online version for this chapter can be found at DOI 10.1007/978-3-319-47471-7_1

J.-M. Josselin (✉)
CREM-CNRS, University of Rennes, Rennes, France
e-mail: jean-michel.josselin@univ-rennes1.fr

A. Marciano
LAMETA-CNRS, University of Montpellier, Montpellier, France
e-mail: alain.marciano@umontpellier.fr

G.B. Ramello
Università del Piemonte Orientale, Alessandria, Italy
e-mail: giovanni.ramello@uniupo.it

A. Marciano and G.B. Ramello (eds.), *Law and Economics in Europe and the U.S.*, The European Heritage in Economics and the Social Sciences,
DOI 10.1007/978-3-319-47471-7_13

In P. 14, reference “An Economic Theory of Clubs. *Economica* 32: 1–14” for Buchanan, J.M. 1964 is updated as “What Should Economists Do? *Southern Economic Journal* 30: 213–222.” and in the next reference the author name “Buchanan, James M. 1958” is changed to “Buchanan, J.M. 1958”.

In P. 14, the references are rearranged and it should read in the order as Buchanan, J.M. 1958; Buchanan, J.M. 1959; Buchanan, J.M. 1964; Buchanan, J.M. 2007.