Probabilistic Study of Voting Rules: A Tale of Two Volumes

Hannu Nurmi

1 Introduction

I would like to begin with a brief personal note on my long scholarly cooperation and friendship with the honoree of this book, Manfred J. Holler. I have known Manfred for more than 40 years. What got us together was an interest in public choice and applied game theory, in those days not widely studied outside the United States. For economists, the public choice was apparently too 'political'. For political scientists, in turn, the game theory seemed too simplistic and the associated methodological individualism downright irrelevant. So, it took a hefty dose of professional courage for a young scholar to become an advocate and, indeed, a standard-bearer of public choice and applied game theory. And yet, Manfred soon became one. In the late 1970s, he founded *Munich Social Science Review* and a few years later *European Journal of Political Economy*, both journals of clear game theory and public choice emphasis. In the early 1980s, he edited one of the classics on applications of game theory to politics: *Voting, Power and Voting Power* Holler (1982). The study of voting power has been—and continues to be—one of the central themes in his scholarship. As a contributor to these fora, I shared much of Manfred's enthusiasm in these areas. He devised the public good index around the same time Holler (1978). Over the past decades, it has been widely and intensively debated and applied, most recently by Manfred and one of the editors of this volume Holler and Rupp (2020). I am mentioning these early works of Manfred because we seemed to be tiptoeing around voting procedures and probability models without actually combining the two as has been done in the two volumes discussed in this paper; one of these volumes appeared only a few years before our cooperation began. Truly grateful for the past decades

The author thanks Benoît Le Maux for constructive comments on an earlier version of this paper.

H. Nurmi (✉)
Department of Philosophy, Contemporary History and Political Science, University of Turku, Turku, Finland
e-mail: hnurmi@utu.fi

M. A. Leroch and F. Rupp (eds.), *Power and Responsibility*,
https://doi.org/10.1007/978-3-031-23015-8_11

of friendship and cooperation, I very much hope that the following pages will be of interest to the honoree of this book.[1]

Voting is about making a choice, either of a person, of a group of persons or of a policy. No doubt choices are made in other ways as well, e.g. by bargaining, by lot, by delegation or by rules of succession. Yet, voting is often regarded as the most democratic way of making collective choices. Indeed, voting is typically regarded as a necessary, albeit not sufficient, condition of democratic governance. Over centuries, a large number of different voting systems have been devised for apparently the same purpose: to tease out the will of the voters. Since the systems are different, it follows that they at least occasionally result in different outcomes when processing the same voter input. This observation motivates the study of properties of various voting systems. At the same time, the question arises as to what constitutes a plausible or reasonable collective choice under various constellations of voter opinions.

This paper discusses a particular, viz. probabilistic, approach to the study of voting rules. It complements the evaluations that focus on norms, conditions or criteria that may or may not characterize the rules. The best-known results of social choice theory establish compatibilities or incompatibilities of those characteristics. Arrow's (1963) impossibility theorem on preference aggregation norms, Sen's (1970) impossibility result on Paretian liberal in resource allocation, Moulin's (1988) result on the incompatibility of participation and Condorcet's principle, as well as Gibbard's (1973) and Satterthwaite's (1975) theorem on general manipulability of social decision functions, are examples from this genre.

Alongside these relatively high-profile works, another approach to voting rules has developed, viz. the probabilistic modeling of voting systems. The main focus of these models is on the probability of encountering various kinds of anomalies and other peculiarities that are related to voting rules under specific circumstances. How likely are we to encounter such circumstances in practice? What kinds of factors contribute to the emergence of such circumstances? How different in the end are various voting rules in terms of the ensuing outcomes? In fact, probabilistic notions were present already in Condorcet's *magnum opus* Black (1958); Condorcet (1785). It must be said, though, that the early reception of these ideas was not uniformly enthusiastic. To wit, some 80 years after its publication, the probability aspect of Condorcet's work received a scorching criticism from Isaac Todhunter (Todhunter, 1865, p. 352):

> We must state at once that Condorcet's work is excessively difficult; the difficulty does not lie in the mathematical investigations, but in the expressions which are employed to introduce these investigations and to state their results: it is in many cases almost impossible to discover what Condorcet means to say. The obscurity and self contradiction are without any parallel, so far as our experience of mathematical works extends; some examples will be given in the course of our analysis, but no amount of examples can convey an adequate impression of the extent of these evils. We believe that the work has been very little studied, for we have not observed any recognition of the repulsive peculiarities by which it is so undesirably distinguished.

[1] Of course, this very brief exposition is not intended as a listing of Manfred's scholarly contributions, not even the most influential ones, but as a background to the following pages.

Indeed, had it not been for Nanson's and Black's rediscoveries, Condorcet's work on voting might well have landed in oblivion Black (1958); Nanson (1883). However, Condorcet's contributions that today bear his name are largely independent of his probability calculus. As the evolution of social choice theory in general, the probabilistic tradition has been highly discontinuous. The first book-length work in the present authors' knowledge did not appear until 1970s. The recent (2021) publication of *Evaluating Voting Systems with Probability Models. Essays by and in Honor of William Gehrlein and Dominique Lepelley* edited by Diss and Merlin (2021) nearly coincides with the 50'th anniversary of the pioneering volume in this genre: *Probability Models of Collective Decision Making* edited by Niemi and Weisberg (1972). This paper aims to provide a brief and non-technical overview of the development of probabilistic voting research in the light of the two volumes just mentioned.

Unsurprisingly, no author has contributed to both of these books.[2] What the volumes have in common is that they deal with probabilistic methods in addressing problems related to voting. A glance at the respective tables of content reveals similarities and differences in foci: constitutional design, coalition formation, and spatial models of party competition are predominant topics in the Niemi and Weisberg book, while the distinctive subjects of the slightly more voluminous Diss and Merlin volume are voting paradoxes and manipulability of voting rules. The study of cyclic majorities has an important place in both volumes. It is only natural and, at the same time, a measure of progress that the Diss and Merlin volume gives a more nuanced picture of the paradox of cyclic majorities and the related concept, Condorcet efficiency, i.e. the probability of a Condorcet winning alternative being elected by various voting methods. The progress of the literature on probabilistic modeling is also reflected in the set of chapters of the Diss and Merlin volume focusing on computational techniques, a subject not much known outside computer science in early 1970s.

In a way, probabilistic modeling represents a return to the roots of voting theory since in the late eighteenth century Marquis de Condorcet was preoccupied with jury decision-making under the assumption that the jurors have an individual probability of passing a correct judgment in dichotomous decision settings Black (1958). The task the Marquis set for himself was to find a decision rule that would maximize the probability of the collectivity making correct decisions. His main result—nowadays known as the Condorcet Jury Theorem (CJT)—suggested that if all voters have the same probability of making the correct choice in dichotomous choice situations and if this probability is higher than $1/2$, then the probability of the majority being right in these kinds of choice situations is larger than the individual probability and tends to unity with the increase of the number of voters. In contrast, if the individual probability of being right is smaller than $1/2$, the probability that the majority is correct is smaller than the individual one and tends to zero as the number of voters grows without limit. The CJT approach is still being pursued today, but is—

[2] While not an author in the Diss and Merlin volume, Peter C. Fishburn—one of the contributors to the Niemi and Weisberg book—certainly had an important role as the supervisor and mentor of William V. Gehrlein, a contributor to and one of the two honorees of Diss and Merlin (2021).

somewhat surprisingly—represented in neither the Diss and Merlin nor in the Niemi and Weisberg volume. Rather the starting point is the standard one of each individual being endowed with complete and transitive preference relations (rankings) over the alternatives. Probability enters the picture via assumptions about the distributions of voters over the preference of rankings.

2 The Principles of Probability Modeling of Voting Rules

As all scholarly activity, probabilistic modeling can be characterized by its goals and methods for achieving those goals. Probability models aim at studying the outcomes ensuing from various voting rules in specific environments. The outcomes depend on the preference profiles of candidates as well as on the way preferences are related to voting strategies. Very often the focus is on reported preferences, i.e. the voting strategies rather than on the underlying 'true' preferences. Thus, the issues related to sincere vs. sophisticated voting are glossed over. This is also the case in the contributions contained in the earlier of the two volumes discussed here. The study of outcomes resulting from voting rules may reveal incompatibilities between various choice desiderata or some intuitively bizarre occurrences. Some of the latter are known as paradoxes. The specific contribution of the probability models is to address the question of how common these occurrences are in various environments, i.e. in various types of preference profiles. It is, however, also possible to consider the likelihood of benefiting from sophisticated voting in various profile types. Regarding the role of probability models in the study of voting rules Gehrlein and Lepelley (2004, p. 141) conclude that:

> Generally speaking, impossibility theorems are essentially qualitative: they state that some paradoxes or difficulties are possibilities but leave open the likelihood of their occurrences. Thus, in the absence of empirical data, probability calculations have proved to be useful.

Another contribution of probability models in these authors' view is in the comparison of voting rules that in principle all fail on a specific criterion or desideratum. The models can then provide important information regarding their likelihood of failure in specific environments. The procedure is straightforward in principle. First, generate all profiles of the desired number of voters over the desired set of candidates. Second, tally the number of profiles where the procedure under study violates the desideratum of interest (e.g. yields a cyclic collective preference relation, elects a Condorcet loser or leads to a voting paradox of some sort). If all profiles are assumed to be equally probable, the tally divided by the number of possible profiles gives the probability of desideratum violations. The population of profiles from which the specific profile sets are drawn (e.g. those exhibiting majority cycles or single-peaked preferences) can simply be generated by sampling one ranking at a time or obtained by constructing probability representations. The latter is typically used in obtaining analytic formulae for relevant probabilities (e.g. of violating the Condorcet principle or choosing the Condorcet loser or ending up with different winner sets).

Two types of profiles stand out in the literature: (i) impartial culture (IC) profiles, and (ii) impartial anonymous culture (IAC) profiles. In the former, the individual preference rankings are randomly and independently drawn (with replacement) from the population of all possible rankings so that any conceivable ranking is equally probable. In the latter, all conceivable *distributions* over the preference relations are deemed equiprobable. Both cultures are intuitively pretty far from describing any real-world electorate, but (ii) is more commonly used. While both (i) and (ii) have very limited descriptive value, they provide a useful benchmark for comparative evaluations of voting rules in terms of various norms and criteria.

3 Common Themes in the Two Volumes

It is only natural that in an active research field topics of interest change with the passage of time. Hence, one can expect to find many new topics in Diss and Merlin (2021) that were not dealt with in Niemi and Weisberg (1972). Similarly, many themes focused in the latter have been omitted in the former volume.

3.1 Cyclic Majorities

One of the common themes in Niemi and Weisberg (1972) and Diss and Merlin (2021) pertains to concepts related to the name of Condorcet. However, the concept of Condorcet winner or loser was not yet in standard use in the early 1970s and not directly applied in the former volume, but the notion of majority cycle was well-known and a focus of interest in the Niemi and Weisberg volume. Bowen discusses the possibility of cyclic majority preference relation underlying the votes in the U.S. Senate (Bowen 1972). It is well-known that the U.S. Senate resorts to the amendment procedure which is based on a pairwise comparison of decision alternatives with the winning alternative surviving in each comparison to face the next one in accordance with the agenda. When all alternatives have been present in at least one pairwise comparison, the winner of the last comparison is declared the overall winner. With k alternatives one thus conducts $k - 1$ comparisons. This obviously falls dramatically short of the $k \times (k - 1)/2$ comparisons required to directly establish the occurrence or non-occurrence of a cyclic majority. As the voting records only reveal the voting behavior of each senator in various pairwise votes, it is not possible to ascertain all the respective individual preferences relations. Suppose we have three policy options: A, B, and C. If the agenda dictates that the first vote be taken between B and C and the second vote between the winner of this vote and A, we cannot with certainty infer the full ranking of a voter who votes for B in the first and—assuming that it is B that faces A in the second ballot—in the second ballot. We do know that this voter (if voting sincerely) prefers B to both A and C, but we do not know the voter's preference between A and C. Had this voter instead voted for A in the second ballot, we could infer that the voter's ranking is ABC. Suppose that A represents the *status quo*, i.e. no change, C a legislative proposal and B an amendment to C. Since three alternatives call for two votes, we can denote any voter's voting strategy by an ordered pair (x, y) where $x, y \in \{yes, no\}$ and where the first element indicates whether the voter prefers the amendment to the original motion ('yes' if he/she does, 'no' if he/she doesn't), and the second element indicates whether the voter prefers the winner of the first ballot to the *status quo*. Let $n(x, y)$ be the number of voters with the voting strategy (x, y). Bowen derives four necessary conditions for the emergence of what he calls the voting paradox (cyclic majority relation):

1. the outcome is A (i.e. *status quo*)
2. $n(no, yes) + n(no, no) > n(yes, yes) + n(yes, no)$
3. $n(yes, no) + n(no, no) > n(yes, yes) + n(no, yes)$
4. combining the two last-mentioned conditions yields $n(no, no) > n(yes, yes)$.

The article proceeds to analyze 111 bills subjected to the roll call vote in the Senate. With probabilistic assumptions applied to the preference relations not directly derivable from the voting records, he concludes that only two bills were associated with a high probability of the voting paradox.

The same problem and approach are pursued by Weisberg and Niemi in the same volume (Weisberg and Niemi 1972). Their findings deviate somewhat from those of Bowen either due to the modeling of the voting procedure or due to the assignment of preferences of some voter groups whose preferences are not directly observable. Perhaps more importantly, Bowen classifies as voting paradoxes those cases where the winning alternative is not a majority (i.e. Condorcet) winner, while Weisberg and Niemi (1972, fn 17) look for vote sequences where the winning alternative(s) would have—in the light of the estimated preference rankings—been defeated by at least one other alternative. The difference can be illustrated by the profile of Table 1 (Nurmi, 1999, p. 28). In this 8-voter profile, both A and B beat C (C being the Condorcet loser), but neither beats the other. Thus, there is no Condorcet winner. At the same time, there is no Condorcet cycle. Thus, Bowen would count profiles like Table 1 as instances of the voting paradox, while Weisberg and Niemi would not as neither A nor B are defeated by any other alternative.

The probability of the voting paradox is also the subject of Bjurulf's (1972) article where the main focus is not so much on the specific probability estimates of encountering a cyclic majority but rather on the way voting blocs affect the paradox probability. In other words, Bjurulf studies the effect of introducing specific types of

Table 1 No Condorcet winner and no Condorcet paradox

3 voters	3 voters	1 voter	1 voter
A	B	C	C
B	A	A	B
C	C	B	A

dependence between voters' opinions on cycle probability. His work uses computer simulations to generate electorates with a small number of blocks (parties) in order to find out the effect of relative sizes of the blocks (equal or unequal) and the degree of conflict between the blocks on the probability of majority cycles.

Shepsle (1972) discusses the significance of impreciseness of political positions in the formation of collective preference relations and their acyclicity. He suggests that considering decision alternatives as risky prospects may explain why expected utility maximizing voters with cycle-generating preferences over certain alternatives may avoid majority cycles over the corresponding risky prospects. Thus, making the alternatives more realistic in the sense of allowing for some impreciseness in the voting setting may do away with the cyclic majority in terms of precise positions. Furthermore, the rules of entry of decision alternatives may explain why proposals involving new policy dimensions are often rejected (e.g. by the chairperson) as non-germane. Thus, one source of voting paradox may play a smaller role in practice than in theory.

In the Diss and Merlin volume, Condorcet's paradox is discussed under the topic of Condorcet efficiency of voting rules. Moreover, the paradox is treated as one of several counterintuitive occurrences potentially encountered in the application of voting procedures. The Condorcet efficiency of a voting rule denotes the probability or—as the case may be—the relative frequency of the rule ending up with the Condorcet winner in the Condorcet domain, i.e. in a set of profiles where such a winner exists. So, for example, if procedure P elects the Condorcet winner in eight profiles out of 100 profiles with a Condorcet winner, its Condorcet efficiency is 0.08.

The Diss and Merlin volume begins with Gehrlein and Lepelley's article (2021) Condorcet efficiency of some common (such as the plurality rule, Borda count, plurality elimination, and approval voting) and some fairly uncommon voting rules (such as negative plurality, negative plurality elimination, and Borda elimination) under IC's and IAC's. As stated above, in the former cultures, each voter's preference ranking is drawn randomly and independently from the set of all possible rankings. In other words, for each voter, each preference ranking is equally probable. In IAC, in turn, all distributions of voters over preference rankings are deemed equally probable. Most of what is known about the Condorcet efficiency is based on these authors' earlier works. The new results of this chapter pertain to the effect of potential abstentions of voters to the Condorcet efficiency of various voting rules. This effect turns out to be significant. As in the no-abstention models, the dependency between the voters seems to play a crucial role in determining the outcomes and, in particular, the Condorcet efficiency of voting rules.

The effect of abstentions is also discussed in the chapter by Diss et al. (2021a). More specifically, these authors focus on weighted scoring rules and their Condorcet efficiency. These rules assign scores to various positions in preference rankings and determine the collective rankings in terms of the score sums of alternatives. In addition to abstention, the authors also consider the effect of indifference in preference relations. The main results pertain to the limit values of the Condorcet efficiency when the number of voters is increased. Surprisingly, the Borda count does not turn out to maximize the Condorcet efficiency, contrary to what has been shown in many earlier studies that do not consider the possibility of voter abstention.

The article of Diss et al. (2021b) addresses a topic that has received relatively scant scholarly attention, viz. the effect that the "closeness" of the election has on the Condorcet efficiency of various voting rules in large electorates.[3] The closeness here refers to the rule-independent opinion distributions preceding the actual voting. Intuitively, one would expect that the presence of one formidable candidate would increase the Condorcet efficiency of rules that elect strong Condorcet (a.k.a. absolute) winners, such as plurality, plurality runoff, instant runoff or Bucklin's rules. Hence, in close elections, a wild guess would be that the effect is the opposite: the closer, the less Condorcet efficient. Using a specific index of election closeness, the authors show that this wild guess is, indeed, a guess: it holds for some rules, but not for all. A notable exception is the Borda count under which the (high) Condorcet efficiency seems to remain largely unaffected by the closeness. It should be added, though, that the results hold for three-candidate elections only.

Brandt et al. (2021) deal with two paradoxes: the Condorcet loser paradox (the eventual election of a Condorcet loser) and the agenda contraction paradox. The former occurs when a preference profile is found or (more commonly) constructed so that the voting rule under investigation results in an alternative that would be majority defeated in pairwise contests by every other alternative. The latter paradox occurs when it turns out that, in a given profile, the rule under study would result in a given alternative, but if some alternatives other than the given one were removed from the profile, another alternative would emerge as the winner.[4] Brandt et al. introduce a (in voting contexts) new methodology, Ehrhart theory, for the analysis of these paradoxes in four-alternative situations mainly under the IAC assumption. The article concludes that the probability of encountering an instance of the Condorcet loser paradox when resorting to some well-known Condorcet extensions (MaxiMin, Dodgson, Tideman, Young) is so low as to make it practically irrelevant under the cultures studied. So, in the authors' view the vulnerability to the Condorcet loser paradox cannot be used as grounds for dismissing those extensions. On the other hand, the occurrence of the agenda truncation paradox seems far more likely.

The picture emerging from the above chapters in Diss and Merlin (2021) is much more nuanced than the one outlined in the chapters of Niemi and Weisberg (1972) touched upon above. At the same time, Diss and Merlin (2021) provides a wider setting for the study of paradox of cyclic majorities. Indeed, Gehrlein and Merlin (2021) analyze the probability of Ostrogorski's paradox, while Belayadi and Mbih (2021) focus on the probability of reversal symmetry violations under various voting rules. The former paradox is related to aggregation of entries in a $k \times m$ matrix of a's and f's each entry indicating, say, that the candidate or applicant for a job qualifies (does not qualify, respectively) on criterion i in evaluator j's opinion with $i = 1, \ldots, k$ and $j = 1, \ldots, m$. In this case, the entry $(i, j) = a$ ($(i, j) = f$, respectively). Sup-

[3] This approach differs from (in fact, reverses) the one where the effects of voting rules on "closeness" of results are sometimes analyzed.

[4] This paradox is also known as the violation of the Chernoff or heritage property.

Table 2 Strong Ostrogorski's paradox

	Criterion 1	Criterion 2	Criterion 3	Aggr.value
Evaluator 1	f	f	a	f
Evaluator 2	f	a	f	f
Evaluator 3	a	f	f	f
Evaluator 4	a	a	a	a
Evaluator 5	a	a	a	a
Aggr.value	a	a	a	a or f

posing that each evaluator considers each criterion equally important, it makes sense to summarize each evaluator's opinions as either a if his/her evaluations of the candidate have more a- values than f-values and vice versa. Considering all evaluators we thus have m summary entries each either a or f. Using the same principle, we can now derive an aggregate evaluation of a candidate by assigning the candidate the value a (f, respectively) if there are more a's(f's) in the m columns. This would amount to columns-first aggregation.

Similarly, one could resort to the rows-first aggregation and derive the aggregate value for the k criteria. Now it may happen that one ends up with contradictory overall evaluation so that the columns-first aggregation leads to value a and the rows-first to f or vice versa. Should this happen, we have an instance of Ostrogorski's paradox. If one of the conflicting evaluations results from 'unanimous' aggregate evaluations (either all evaluators give the same aggregate assessment of the candidate or all criteria suggest the same aggregate value for the candidate), then we have an instance of the strong Ostrogorski's paradox. Table 2 where $k = 5$ and $m = 3$ illustrates.

Here a job applicant is being evaluated by five evaluators, each assigning him either 'pass' (denoted by 'a') of 'fail' (denoted by 'f') on each criterion. The overall mark of the applicant given by the evaluator is obtained as the value that most often occurs in this evaluator's criterion-wise evaluations of the candidate. Similarly, the criterion-wise mark of the applicant is obtained as the value given by the majority of the evaluators on this criterion. These values are presented in the right-most column and lower-most row in Table 2. The paradox consists in the observation that the two ways of aggregating values lead to conflicting results: rows-first yields 'f', while columns-first gives 'a'. Note that the latter conclusion stems from, not just from a majority, but from all aggregated criterion-wise values suggesting it. In fact, 60% of the voters vote for 'a' on each criterion and, yet, 60% of voters support 'f' on the basis of their aggregated evaluations. Gehrlein and Merlin find that the strong version of Ostrogorski's paradox is extremely unlikely to happen under the IC assumption.

Reversal symmetry paradox (a.k.a. preference inversion paradox or reversal bias) occurs when, under a given voting rule, the reversal of a preference profile leads to the same outcome as the outcome ensuing from the application of the rule in the original profile. Table 3 gives an instance of the paradox under the plurality runoff

Table 3 Reversal symmetry paradox and plurality runoff

4 voters	3 voters	2 voters
A	B	C
B	C	A
C	A	B

system.[5] In the initial profile, there is a runoff between A and B, whereupon A wins. In the profile where all rankings are inverted, we get a runoff between A and C where A again comes out victorious. Hence, the same outcome ensues in a profile and its complete inverse.

Belayadi et al. derive closed formulae for the probability of reversal symmetry paradox in IC's as well as in IAC's for selected voting rules for a small number of alternatives (3 and 4). It seems that the results largely support the contention of Regenwetter et al. (2006) according to which IC tends to exaggerate paradox probabilities. It should be born in mind, though, that while the IAC seems in general to be associated with lower paradox probabilities than the IC, the finding doesn't say much about the likelihood of encountering the said paradoxes in real world voting situations. We shall return to this point later on.

3.2 The U.S. Political Institutions

In the 1970s, the probabilistic modeling was largely conducted in the U.S. universities. This is reflected in the author list of Niemi and Weisberg (1972): only two authors out of 17 had a non-American affiliation. In the Niemi and Weisberg volume two U.S. institutions—the Congress and the Supreme Court—are singled out. Bowen's contribution was already discussed above. Its main focus is in the possibility of cyclic majorities underlying the votes in the Senate. Koehler (1972) examines the coalition formation in the U.S. House of Representatives using as the benchmark the size principle, i.e. the hypothesis that coalitions of minimal winning size will form. In the words of Riker (1962, pp. 32–33):

> In social situations similar to n-person, zero-sum games with side-payments, participants create coalitions just as large as they believe will ensure winning and no larger.

Using the Brams-O'Leary index of voting agreement (Brams and O'Leary 1970), Koehler examines the roll call votes in eleven sessions of the House of Representatives to assess the predictive value of the size principle. Both party and individual level agreement are calculated. The concept of probability enters the discussion via the Brams-O'Leary index which is based on expected numbers of voting agreement. Rohde's focus is on the validity of the size principle, but in a context where zero-sum assumption underlying Riker's construction seems at first blush somewhat bizarre, viz. in the U.S. Supreme Court decision-making (Rohde 1972). What the size principle would then imply in the nine-member Court is that a five-member opinion

[5] The profile is a minor simplification of the one presented in (Felsenthal and Nurmi, 2018, p. 34).

Table 4 An instance of the referendum paradox (Nurmi, 1999, p. 77)

Ref. stand	MP1	…	MP167	MP168	…	MP200	Row sum
Yes	7.000	…	7.000	15.000	…	15.000	1.664.000
No	8.000	…	8.000	0	…	0	1.336.000

coalitions be formed. However, Rohde argues that in decisions under an external threat to the Court, the coalitions will be larger than minimal winning. Assuming that each judge votes 'yes' or 'no' with probability $1/2$, Rohde computes the theoretical probabilities that majority coalitions of various sizes form. He then compares empirically observed coalition data with these values. He finds that there is a significant difference in the probabilities predicted by the size principle and the empirical data. So, the principle doesn't seem applicable to the Supreme Court decision-making either in the case of opinion situations or under external threat. He concludes that a more nuanced study of various decisions situations is called for in order to find out the domains where the size principle is applicable as a predictor.

Now, some 50 years later, the probabilistic study of voting rules has a different focus when it comes to political institutions. Instead of the study of roll call votes in the main U.S. federal institutions, the idea of federalism itself is brought to the agenda of study in the Diss and Merlin volume. Specifically, two multi-author contributions focus on the representative principles in federations. Feix et al. (2021) study two-tier systems of governance where on one tier the voters elect representatives for an assembly that constitutes the second tier. The ultimate decision maker is the latter tier. Numerous institutions are instances of the two-tier arrangement: the Council of Minister of the European Union and the Electoral College of the U.S. being perhaps the best-known ones, but basically all representative institutions are two-tier systems: the voters elect parliamentarians (or other representatives) to make binding decisions on legislation, policies, allocation of resources, etc. Because the choices (decisions) made on the second (representative) tier are not ones voted upon on the first tier and yet the voters on the latter often have opinions regarding those choices as well, it may happen that the choices ensuing from the decision making of the representatives would be voted down by the electorate at large had it been given the opportunity to vote directly on those choices. This kind of occurrence is sometimes called the referendum paradox. A fictitious example from Nurmi (1999) illustrates. A parliament consists of 200 members (MP's) each elected from a single-member constituency. The voting population is assumed to be 3 million. This is distributed into 200 constituencies of equal size, viz. 15000 voters. The country ponders upon entering an economic union. The stands of the voters in each constituency are presented in Table 4. Should a referendum be arranged, 'yes' would clearly be the victorious stand with more than 55% of the voters voting for it. However, should the issue be decided in the parliament, a whopping $5/6$ of the MP's would 'in good faith' support 'no' correctly convinced that they are expressing the will of the majority of their constituency.

Feix et al. evaluate various voting rules in terms of a criterion they have devised, viz. majority efficiency. A method is the more majority efficient the less likely it makes the instances of referendum paradox to arise. The article discusses the Penrose–Banzhaf (independent population) probability model (Penrose 1946; Straffin 1977) and compares it with May's (homogeneous population) model (May 1948; Straffin 1977) to find out whether the allocation of seats to states in accordance with Penrose's proposal, i.e. in proportion to the square-roots of the respective populations, leads to a more majority efficient allocation than May's model that suggests seat allocation in direct proportion to the populations. The authors resort mainly to Monte Carlo simulations to estimate the majority efficiencies of various allocation rules. They also study rules that assign seats to states according to the formula

$$a_i = n_i^{\delta}$$

where a_i is the number of seats allocated to state i, n_i the population of state i and δ a coefficient ranging from 0 (one seat per state) via $1/2$ (the square-root rule) and 1 (strict proportionality) to ∞ (all seats to the most populous state). For each situation that includes a fixed number of states and decision alternatives, the authors run 10^5 or 10^6 simulated elections under IC and IAC assumptions to find out the majority efficiencies of various rules. It turns out, e.g. that the square-root rule is less obvious winner in IC cultures than has been thought.

The article of De Mouzon et al. (2021) can be considered a companion article of Feix et al. (2021) in the sense that both of these articles make the same probabilistic assumptions regarding the electorate. The central concept studied is the justice of the election procedure in two-tier systems. This contribution utilizes the standard power indices—the Banzhaf and Shapley–Shubik indices—and looks at the distributions of these values within each state in order to judge the justice of the system. A whopping 10^{12} simulation rounds were conducted in each state of the U.S. to obtain estimates of pivot probabilities of voters in each state. Owen's study from 1970's is used as a benchmark in comparing deviations from perfectly equal pivot probabilities between the states Owen (1975). It turns out that the main conclusions regarding electoral justice hinge on the probabilistic population models. Specifically, May's model leads to the conclusion that violations of electoral justice mainly take place at the expense of populous states, while both Banzhaf and Shapley–Shubik probability models suggest the opposite, i.e. that the deviations occur mainly to the disadvantage of small states.

4 Non-overlapping Themes

The preceding section singles out a couple of topics discussed in the two volumes reviewed here. Apart from these, there are several themes that are discussed in one of the books only. One of these is coalition formation and another spatial political competition. This doesn't mean that study of these areas has stopped over the 50 years that followed the publication of the Niemi and Weisberg volume. Coalition formation remains an important topic in basically all multi-party democracies. Moreover, Diss and Merlin (2021) has, as its title suggest, a probabilistic emphasis, while most political coalition theories are deterministic. Nonetheless, one cannot escape the impression that the probabilistic study of voting rules is today mainly in the domain of other branches of science than political science.[6]

The Diss and Merlin book contains a host of topics not touched upon in the Niemi and Weisberg one. Given the time lapse between the two books, this is only natural. Much of the work done on voting rules is based on problems left open by earlier studies. In fact, this is the way many new voting rules are invented. However, many open problems can be addressed with methods borrowed from disciplines not directly related to voting. Probability modeling is, of course, a good example. However, in addition to probability theory and statistics, computer science has entered the field in a big way creating a new sub-field in voting theory: computational social choice. This development has its origins much later than the Niemi and Weisberg volume. As probability modeling, the computational social choice has to confront the question of relevance: what can be achieved through the use of the tools and techniques stemming from computer science? In the case of probabilistic modeling the answer was to give a more nuanced picture of the virtues and vices of voting rules. Instead of just[7] stating that a procedure fails on a criterion of performance, we need to have some idea about how likely such failures are. Probability models provide—among other things—relevant information insofar as they give estimates about the failure probabilities and—perhaps more importantly—about factors increasing or decreasing the failure probabilities. Similarly, the computational techniques provide information about the practical difficulty of finding the winners or about the difficulty of successfully misrepresenting one's preferences and similar aspects affecting the relevance of the voting theory results. In a similar vein, Brandt et al. (2021) apply the computational approach of Ehrhart (1962) to find out the likelihood of the no-show paradox. Also, Karpov (2021) provides an overview of combinatorial methods in analyzing voting rules. Moyouwou et al. (2021) present an overview of the pioneering Fishburn–Gehrlein method for computing the number of integer solutions to a set of linear inequalities, a method that has subsequently been often applied, notably by Gehrlein and his associates, in probabilistic voting studies. The authors also discuss Ehrhart's conjecture and provide outlines for associated algorithms. Overall, the menu of techniques is thus considerably more extensive than in the Niemi and Weisberg book.

When it comes to topics not present in the earlier volume, two parts of Diss and Merlin (2021) stand out: one on resistance to manipulations and the other on game theory. An obvious reason for the absence of the former theme in the Niemi and

[6] This is not the proper place to address the age-old topic of what constitutes a science *sensu stricto*. No doubt voices have occasionally been raised to argue that 'science' should denote natural sciences and should not be applied to the study of such 'soft' entities as politics—or economics for that matter.

[7] This expression is not intended to play down the significance of findings regarding paradoxes or incompatibilities between performance criteria. Such findings may be exceedingly difficult to make.

Weisberg volume is that the idea of strategic voting and the associated outcome manipulability had not been widely studied at the time of the publication of the volume.[8] The part on game theory focuses on rather specific topics in game theory rather than on the general foundations of voting games. The Diss and Merlin volume concludes with a progress report on probability calculations based on the IAC assumption. This assumption was not used in the Niemi and Weisberg volume where the dominant assumption is IC.

5 Progressive Program

All this leads to the conclusion that the probability modeling of voting rules is what used to be called a research program in the sense of Imre Lakatos: a theoretical-empirical system with a distinctive set of core assumptions not challenged within the program and surrounded by lower-level theoretical results challenged occasionally with empirical evidence Lakatos (1977). Should serious discrepancies between the obtained evidence and theoretical statements or principles ensue, those lower-level statements are dispensed with or replaced with others, while the core tenets are maintained. The probability modeling tradition fits these hallmarks of a research program. It is also a progressive entity: in time, the theoretical foundation gets more nuanced and the range of procedures and their workings expand so that the new structures embrace the older ones in accounting for the successes of the latter while correcting their failures. In the probability tradition, the core beliefs pertain to the basic constituents: the culture assumptions, the individual preferences, the voting strategies, and the outcomes.

The basic finding in the comparison of the two volumes is that the models and research strategies have become far more nuanced and sophisticated with new insights from other disciplines, mathematics and computer science, in particular. Progress is also visible in the developments of computational capacity of modern machinery. What in the 1970s seemed like an adequate number of (simulated) elections to guarantee reasonable convergence of the estimated probabilities, say about 10^4 rounds of elections, required a considerable amount of computational resources limiting the simulation modeling to relatively small candidate and voter sets. Nowadays, these restrictions have been relaxed to a substantial degree. Importantly, also the algorithms used in solving various problems through simulations have become more versatile. In short, the probability modeling has made progress in theoretical, methodological, and technical aspects. At the same time, the small community of scholars contributing to the Niemi and Weisberg volume has grown to a vast network of researchers from all continents.

---

[8] See, however, the pioneering work of Farquharson (1956) where the idea of sophisticated voting as a form of successful preference misrepresentation and the concept of straightforward voting rule as a rule immune to successful misrepresentation was introduced. These ideas were further developed in Farquharson (1969). See also Dummett and Farquharson (1961).

6 What Do We Learn from Probability Models?

In what way do the probability models of voting rules then help us in understanding voting institutions? A partial answer to this is that these models stand midway between the compatibility/incompatibility results of the social choice theory and empirical data on voting results. This is nicely expressed by the two honorees of the Diss and Merlin volume in the quote of Sect. 2 above. The probability models transform basically dichotomous evaluations into graded ones. To put it bluntly, they tell us which ones of bad procedures are less bad than others in terms of specific criteria. This sounds somewhat akin to Tolstoy's famous dictum:

> Happy families are all alike; every unhappy family is unhappy in its own way.

Translating this to the language of voting rules, the happy ones would all satisfy a criterion of goodness. For a voting rule being unhappy would then mean that it fails on a theoretical criterion and being unhappy in its own way would mean that it has its own probability of failing. So, the picture emerging from probability estimates is more precise than the one resulting from studies on the compatibility of norms. From a practical point of view, various 'cultures', e.g. IAC, are profile restrictions and, as Arrow famously showed, restrictions make a great deal of difference in the compatibilities of norms. Accordingly, to rely on estimates based on IC amounts to using irrelevant information in institution design in situations where it is known that some fixed restrictions always hold. Thus, for example, the domain of single-peaked preferences or, more generally, the Condorcet domain is of particular interest in cases where the preferences tend to obey this restriction. The difficulty is that the set of potentially relevant sub-domains is not known beforehand.

The upshot is that the probability estimates of paradoxes do not say much about the likelihood of encountering those paradoxes in practice. After all, we are dealing with statistical inference with at best confidence degrees and intervals. So, the judgments regarding voting rules are bound to be of the type: 'rule A leads to a smaller failure probability on criterion X than rule B assuming that the relevant voter set is drawn randomly from a population of profiles satisfying assumption Y'. This leaves open the possibility that the priority of A *vis-à-vis* B is inverted in some proper subset, say Y', of those profiles satisfying Y. This in itself is natural and even obvious, but in cases where Y' is the profile restriction prevailing in practice, we may be misled in our designs based on probabilities based on Y.

7 By Way of Conclusion

The two volumes reviewed above open a rich panorama of the developments in the field of probability modeling of voting rules. Over the 50-year span, the progress has been dramatic in terms of methodologies employed and in terms of the subjects included in the study. Yet, some crucial features of the Niemi and Weisberg volume remain visible in the Diss and Merlin book. What started as a study of the probability of ending up with cyclic majorities in specific voting bodies has over time developed into a technically sophisticated analysis of many other types of paradoxes. The likelihood of encountering profiles with a Condorcet winner has been complemented with the study of Condorcet efficiency of non-Condorcet voting rules. The assumption of sincere voting has been replaced by a more nuanced view of voter motivations. Accordingly, manipulation probabilities have become the focus of a distinct body of literature. The range of institutions studied has widened from committee-type bodies to representative two-tier systems. These are but some of the most prominent developmental strides so far taken on the way to a better understanding of voting rules from the probabilistic perspective.

References

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

Belayadi, R. & Mbih, B. (2021). Violations of reversal symmetry under simple and runoff scoring rules, pp. 137–160 in Diss & Merlin (2021).

Bjurulf, B. (1972). A probabilistic analysis of voting blocks, pp. 232–251 in Niemi & Weisberg (1972).

Black, D. (1958). *Theory of committees and elections*. Cambridge University Press.

Brams, S. J., & O'Leary, M. (1970). An axiomatic model of voting bodies. *American Political Science Review, 65*, 449–470.

Bowen, B. D. (1972). Toward an estimate of the frequency of occurrence of the paradox of voting in U.S. Senate roll call votes, pp. 181–203 in Niemi & Weisberg (1972).

Brandt, F., Geist, C. & Strobel, M. (2021). Analyzing the practical relevance of the Condorcet loser paradox and the agenda contraction paradox, pp. 97–115 in Diss & Merlin (2021).

Brandt, F., Hofbauer, J. & Strobel, M (2021). Exploring the no-show paradox for Condorcet extensions, pp. 251–273 in Diss & Merlin (2021).

De Condorcet, M. (1785). Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. L'Imprimerie Royale.

De Mouzon, O., Laurent, T., Le Breton, M. & Moyouwou, I. (2021). "One man, one vote" part 1: Electoral justice in the U.S. Electoral College: Banzhaf and Shapley/Shubik versus May, pp. 189–227 in Diss & Merlin (2021).

Diss, M., Kamwa, E., Moyouwou, I. & Smaoui, H. (2021a) Condorcet efficiency of general weighted scoring rules under IAC: Indifference and abstention, pp. 55–73 in Diss & Merlin (2021).

Diss, M. & Merlin, V. (Eds.). (2021). *Evaluating voting systems with probability models*. Essays by and in Honor of William Gehrlein and Dominique Lepelley. Springer.

Diss, M., Pérez-Asurmendi, P. & Tlidi, A. (2021b). The effect of closeness on the election of a pairwise majority rule winner, pp. 75–95 in Diss & Merlin (2021).

Dummett, M., & Farquharson, R. (1961). Stability of voting. *Econometrica, 29*, 33–43.

Ehrhart, E. (1962). Sur les polyèdres rationnels homothétiques à *n* dimensions. *Compte-Rendus de l'Academie des Sciences, 254*, 1273–1302.

Farquharson, R. (1956). Straighforwardness in voting procedures. *Oxford Economic Papers, 8*, 80–89.

Farquharson, R. (1969). *Theory of voting*. Yale University Press.

Felsenthal, D. S., & Nurmi, H. (2018). *Voting procedures for electing a single candidate*. Springer.

Feix, M., Lepelley, D., Merlin, V., Rouet, J. - L. & Vidu, L. (2021). Majority efficient representation of the citizens in a federal union, pp. 163–187 in Diss & Merlin (2021).

Gehrlein, W. V., & Lepelley, D. (2004). Probability calculations in voting theory: An overview of recent results. In M. Wiberg (Ed.), *Reasoned choices* (pp. 140–160). The Finnish Political Science Association: Helsinki.

Gehrlein, W.V. & Lepelley, D. (2021). Analyzing the probability of election outcomes with abstentions, pp. 15–53 in Diss & Merlin (2021).

Gehrlein, W. V. & Merlin, V. (2021). On the probability of the Ostrogorski paradox, pp. 119–135 in Diss & Merlin (2021).

Gibbard, A. (1973). Manipulation of voting schemes. *Econometrica, 41*, 587–601.

Holler, M. J. (1978). A priori party power and government formation: Esimerkkinä Suomi. *Munich Social Science Review 1*, 25–41.

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

Holler, M. J. & Rupp, F. (2020). Some strategic decision problems in networks. *Transactions of Computational Collective Intelligence XXXV*, 133–147.

Karpov, A. (2021) Combinatorics of election scores, pp. 347–366 in Diss & Merlin (2021).

Koehler, D. H. (1972). The legislative process and the minimal winning coalition, pp. 149–164 in Niemi & Weisberg (1972).

Lakatos, I. (1977). *The methodology of scientific research programmes: Philosophical papers* (Vol. 1). Cambridge University Press.

May, K. O. (1948). Probabilities of certain election results. *The American Mathematical Monthly, 55*, 203–209.

Moulin, H. (1988). Condorcet's principle implies the no-show paradox. *Journal of Economic Theory, 45*, 53–64.

Moyouwou, I., Andjiga, N. G. & Mbih, B. (2021). From Gehrlein-Fishburn's method on frequency representation to a direct proof of Ehrhart's extended conjecture, pp. 367–398 in Diss & Merlin (2021).

Nanson, E. J. (1883). Methods of election. *Transactions and Proceedings of the Royal Society of Victoria*, art. XIX, 197–240.

Niemi, R. G., & Weisberg, H. F. (Eds.). (1972). *Probability models of collective decision making*. Charles E. Merrill.

Nurmi, H. (1999). *Voting paradoxes and how to deal with them*. Springer Verlag.

Owen, G. (1975). The presidential election game. *American Political Science Review, 69*, 947–953.

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

Regenwetter, M., Grofman, B., Marley, A. A. J., & Tsetlin, I. (2006). *Behavioral social choice: Probabilistic models, statistical inference and applications*. Cambridge University Press.

Riker, W. H. (1962). *The theory of political coalitions*. Yale University Press.

Rohde, D. W. (1972). A theory of the formation of opinion coalitions in the U.S. Supreme Court, pp. 165–178 in Niemi & Weisberg (1972).

Satterthwaite, M. (1975). Strategy-proofness and Arrow's conditions. *Journal of Economic Theory, 10*, 187–217.

Sen, A. K. (1970). The impossibility of a Paretian liberal. *Journal of Political Economy, 78*, 152–157.

Shepsle, K. A. (1972). The paradox of voting and uncertainty, pp. 252–270 in Niemi & Weisberg (1972).

Straffin, P. D. (1977). Homogeneity, independence and power indices. *Public Choice, 30*, 107–118.

Todhunter, I. (1865/1949). *The history of the mathematical theory of probability from the time of Pascal to that of Laplace*. Chelsea Publishing Company.

Weisberg, H. F. & Niemi, R. G. (1972). Probability calculations for cyclical majorities in Congressional voting, pp. 204–233 in Niemi & Weisberg (1972).