Studies in Choice and Welfare

Rudolf Fara
Dennis Leech
Maurice Salles *Editors*

Voting Power and Procedures

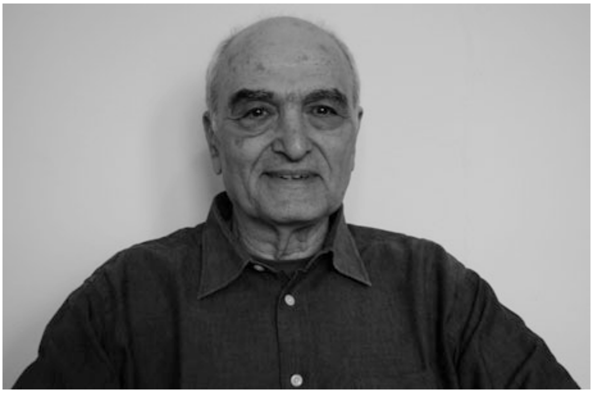
Essays in Honour of Dan Felsenthal and Moshé Machover

Springer

Studies in Choice and Welfare

For further volumes:
http://www.springer.com/series/6869

Dan Felsenthal



Moshé Machover. Photo by Hannah Machover

Rudolf Fara • Dennis Leech • Maurice Salles
Editors

Voting Power and Procedures

Essays in Honour of Dan Felsenthal and Moshé Machover

Springer

Editors
Rudolf Fara
Centre for Philosophy of Natural and Social Science
London School of Economics and Political Science
London
United Kingdom

Dennis Leech
Department of Economics
University of Warwick
Coventry
United Kingdom

Maurice Salles
CREM
Université de Caen Basse-Normandie
Caen
France

Studies in Choice and Welfare ISSN 1614-0311
ISBN 978-3-319-05157-4 ISBN 978-3-319-05158-1 (eBook)
DOI 10.1007/978-3-319-05158-1
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014940732

Printed on acid-free paper

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

Contents

Contributors

Rudolf Fara Centre for Philosophy of Natural and Social Science, London School of Economics and Political Science, London, UK

Dennis Leech Department of Economics, University of Warwick, Coventry, UK

Maurice Salles CREM, Université de Caen-Basse-Normandie, Caen, France

Claus Beisbart Institut für Philosophie, Universität Bern, Bern, Switzerland

Olga Birkmeier Lehrstuhl für Stochastik und ihre Anwendungen, Institut für Mathematik, Universität Augsburg, Augsburg, Germany

Steven Brams Department of Politics, New York University, New York, NY, USA

Sreejith Das Department of Economics, Mathematics and Statistics, Birkbeck College, University of London, London, UK

Abraham Diskin Department of Political Science, The Hebrew University of Jerusalem, Jerusalem, Israel

Maria Ekes Warsaw School of Economics, Warsaw, Poland

Scott Feld Purdue University, West Lafayette, IN, USA

William Gehrlein Department of Business Administration, University of Delaware, Newark, DE, USA

Joseph Godfrey WinSet Group, LLC, Falls Church, VA, USA

Bernard Grofman Department of Political Science, Centre for the Study of Democracy, University of California, Irvine, CA, USA

Manfred Holler University of Hamburg, Hamburg, Germany

Marc Kilgour Department of Mathematics, Wilfrid Laurier University, Waterloo, ON, Canada

Werner Kirsch Fakultät für Mathematik und Informatik, FernUniversität Hagen, Hagen, Germany

Moshe Koppel Department of Computer Science, Bar-Ilan University, Ramat-Gan, Israel

Jessica Langner Fakultät für Mathematik und Informatik, FernUniversität Hagen, Hagen, Germany

Nicola Maaser Universität Bremen, Zentrum für Sozialpolitik, Bremen, Germany

Iain McLean Nuffield College, University of Oxford, Oxford, UK

Nicholas Miller Department of Political Science, University of Maryland Baltimore County (UMBC), Baltimore, MD, USA

Maria Montero School of Economics, University of Nottingham, Nottingham, UK

Sefan Napel University of Bayreuth, Bayreuth, Germany

Hannu Nurmi Public Choice Research Centre, University of Turku, Turku, Finland

Friedrich Pukelsheim Lehrstuhl für Stochastik und ihre Anwendungen, Institut für Mathematik (2. OG, Zi. 3012), Universität Augsburg, Augsburg, Germany

Souvik Roy Economics and Planning Unit, Indian Statistical Institute, New Delhi, India

Dieter Schmidtchen Centre for the Study of Law and Economics, University of Saarland, Saarbruecken, Germany

Wojciech Słomczyński Institute of Mathematics, Jagiellonian University, Kraków, Poland

Frank Steffen Faculty of Law, Business and Economics, University of Bayreuth, Bayreuth, Germany

Department of Economics, Finance and Accounting, The University of Liverpool Management School (ULMS), Liverpool, UK

Bernard Steunenberg Department of Public Administration, Universiteit Leyden, Den Haag, The Netherlands

René van den Brink Department of Econometrics, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Karol Zyczkowski Institute of Physics, Jagiellonian University, Kraków, Poland

Part I
Overview and Interview

Introduction

Rudolf Fara, Dennis Leech, and Maurice Salles

This volume collects the invited essays presented in honour of Dan Felsenthal and Moshé Machover. Most of the papers were delivered at the Voting Power in Practice Symposium, *Voting Power in Social/Political Institutions*: *Typology*, *Measurement*, *Applications* held at the London School of Economics, 20–22 March 2011. The symposium had been planned both to mark the end of 8 years of Leverhulme Trust funding of the LSE's Voting Power & Procedures (VPP) research programme and to celebrate the immense contribution to the field of voting theory by Felsenthal and Machover's (F&M) critically acclaimed monograph *The Measurement of Voting Power* (MVP) published a decade earlier.

The co-celebration was a unique and altogether fitting tribute, to which a brief background sketch will attest, to F&M's landmark book and to the VPP research programme it inspired. The generous research award, in its turn, helped enormously to encourage the prodigious qualitative output from the F&M partnership for more than a decade. MVP was a comprehensive analysis of a priori voting power theory and its measurement, and more. As well as its own important theoretical contributions, it analysed and contextualised the history, and drew on numerous pertinent case studies from the EU, UN and US governance to exemplify the origins and development of this foundational area of social choice.

R. Fara (✉)
Centre for Philosophy of Natural and Social Science, London School of Economics and Political Science, London, UK
e-mail: r.fara@lse.ac.uk

D. Leech
Department of Economics, University of Warwick, Coventry, UK
e-mail: d.leech@warwick.ac.uk

M. Salles
Institute for Social Choice and Welfare, University of Caen, Louvigny, France
e-mail: maurice.salles@unicaen.fr

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare,
DOI 10.1007/978-3-319-05158-1_1,

The book's reviewers at the time were unanimous in their praise: "To say that this book is excellent would be an understatement. It is really remarkable . . . "; "The history of the power indices goes back more than fifty years and is told accurately and completely, for the first time . . . "; "It is at the cutting edge of research in the theory and measurement of a priori voting power, but it is also of practical and political relevance, insofar as it provides a sound basis for the analysis of real-life decision-making processes"; " . . . No one working in the field of formal political theory, institutional design and/or applied social choice theory can afford to ignore it" . . . and so on.

The monograph contributed enormously to a reawakening in the field of voting power; it inspired the creation of the VPP programme that would play a major part in that reawakening and in the further development of voting theory research internationally to the present day. In 2000, just over a year following the book's publication, Machover and Fara founded VPP at LSE's Centre for Philosophy of Natural and Social Science. Dan Felsenthal, Dennis Leech and Maurice Salles joined officially as VPP co-directors the following year, and the late Sir Michael Dummett and Nobel Laureates Kenneth Arrow and Amartya Sen formed the distinguished honorary advisory board to the project soon after. VPP's declared mission proclaimed its dedication to multidisciplinary research in the theory and practice of voting power and procedures with stress laid on the practical application and dissemination of results. The multidisciplinarity and practical application emphasised in the monograph, and featured positively by many of its reviewers, was now welded into the framework of the research vehicle.

In 2001, the Leverhulme Trust awarded VPP a 3-year (extended to 4) Research Interchange Grant to develop an international research network in the field of the measurement of voting power. The focus was on current work in the theory and practice of voting and in its application to the design of international organisations, in particular the system of qualified majority voting in the EU Council of Ministers. The subsequent work by F&M, and by the international research network, particularly at the annual VPP Workshops,[1] expanded significantly on the case study on EUCM voting discussed in the monograph.

In 2007 the Leverhulme Trust funded for 3 years (also extended to 4) a further VPP research network, *Voting Power in Practice*. This new initiative again was developing further themes touched upon in the book, particularly in stressing the methodological importance of bringing the theory to the practitioners. *Voting Power in Practice* focused on the practical application, dissemination and evaluation of research in voting power for improved governance and policy-making. A key objective was to promote mutual understanding and effective dialogue between VPP's extensive international network of voting power theorists from various academic disciplines—economics, political science, mathematics, law and philosophy—and practitioners and their advisers. The emphasis on interaction was

[1] See archived annual workshop proceedings at: www.lse.ac.uk/vpp.

designed to encourage more cross-disciplinary voting power research with a focus on application and policy development.

With a generously funded international research vehicle in place F&M continued their research with originality and vigour, extending their reach productively into most areas of voting theory. Voting procedures were given in-depth treatment at the *Voting Power in Practice* Summer Workshop, *Assessing Alternative Voting Procedures*, held in France from 30 July to 1 August 2010 at the Chateau du Baffy in Normandy.[2] Dan and Moshé made significant individual contributions to the workshop and co-edited its proceedings in a volume published by Springer in 2012, *Electoral Systems: Paradoxes, Assumptions, and Procedures.*

This Festschrift includes 18 invited contributions and opens with a scene-setting interview, reflecting further on the protagonists' insights beyond voting power and its measurement. Although most of the essays, each forming a chapter in the volume, are specifically devoted to voting power rather than to voting procedures, one can observe clearly that voting power analysis cannot be neatly disentangled from the associated voting procedures. We have grouped the chapters under four headings, but our choice of taxonomy should not disguise the very large overlap between these parts.

Part II, *Foundations of Power Measurement*, is devoted to the underpinnings of power measurement. Egalitarian ideas pervade microeconomic theory—through the theory of general competitive equilibrium; social choice theory has its notions of anonymity, equality of decision-makers, neutrality, equal treatment of options, etc. However, from a more commonplace realistic world perspective, inequality prevails. Oligopolies really do exist and there are very rich people and very poor people, weak people and strong people. Power is manifestly everywhere and crucially it is very unequally distributed. Power measurement has undergone major development in the last century. These developments have been very often based on a combinatorial analysis where the underlying probabilistic assumptions are rather rudimentary.

In the chapter "The Measurement of Voting Power as a Special Case of the Measurement of Political Power", Abraham Diskin and Moshe Koppel define political power under which voting power is presented as a special case. The definition of political power they propose is shown to be a generalization of Banzhaf's definition when applied to voting power.

In "On the Measurement of Success and Satisfaction", René van den Brink and Frank Steffen reconsider the notion of satisfaction (related to preference) that is often taken as a synonym of success. For many authors, the notions of power and success are quasi-identical. Several scholars have vindicated this view, in particular Laruelle and Valenciano in several articles and in their book, *Voting and Collective Decision-Making* published in 2008. Van den Brink and Steffen here distinguish between satisfaction and success and show that satisfaction entails success as one component.

[2]See: www.lse.ac.uk/CPNSS/projects/VPP/workshops/8thannualworkshop.aspx for full details of this important discussion of voting procedures.

Sreejith Das introduces a new methodology in the chapter “Voting Power Techniques: What Do They Measure?” that can be employed to calculate any voting power measure, irrespective of the underlying probability model, and to determine what the different indices are really expressing. Power is often defined according to some probability, the power of an individual being associated with the probability that this individual can affect an outcome.

In chapter “Voting Power and Probability”, Claus Beisbart investigates the kind of probability it could be. He provides insights based on philosophical reflections that exceed the strict analysis of voting power to cover themes from equality and rights.

Olga Birkmeier and Friedrich Pukelsheim devote “A Probabilistic “Re-View” on Felsenthal and Machover’s The Measurement of Voting Power” to a general presentation of the probabilistic approach to voting power measurement where voters have three options: vote in favour, vote against or abstain.

Part III, *Power in Two-Tier Voting Systems*, comprises four chapters that consider major questions of multi-state organisations: the two-tier voting systems and the attribution of weights. Two-tier voting systems are standard in representative democracies. Given majority rule, this implies that a party (or a coalition of parties) having more than 50 % of the seats in an assembly has full power. However, in some organisations (the EU for instance), a system must be designed in which assemblies and committees are made of a number of representatives whose power is correlated with the population they represent. In the Electoral College, the body that directly elects the president of the USA, the plurality winner in each state carries all the electoral votes in that state.[3] This sometimes creates paradoxes where the plurality winner at the national level is not the eventual winner.

In the chapter “Square Root Voting System, Optimal Threshold and π”, Karol Życzkowski and Wojciech Słomczyński reconsider the square root rule—that the weight of a representative must be proportional to the square root of the population she represents—and the choice of a quota for a qualified majority—to equalize approximately voting weights and voting power. The co-authors derive an approximate formula for this quota.

In “The Fate of the Square Root Law for Correlated Voting”, Werner Kirsch and Jessica Langner deal with the “democracy deficit”, that is the difference between the outcome of the council vote and the popular vote. To minimize this deficit, the square root rule must be used according to an analysis based on assumptions of voters’ independence. The authors drop the independence assumption and assume a “common belief” among voters.

[3]This is the case in all the states except Maine and Nebraska, each with three electoral votes, in which one electoral vote is awarded to the candidate receiving the most votes in each of the congressional districts, and the remaining two votes are awarded to the candidate that gets the most votes statewide. Although the possibility of a “split” allocation does exist, it has never actually happened.

Nicola Maaser and Stefan Napel also tackle the square root rule of the two-tier voting system in their "The Mean Voter, the Median Voter, and Welfare-Maximizing Voting Weights". In this work, the preference of each representative is supposed to coincide with the preference of her constituency's median voter. They argue that the objective is to maximise the total expected utility generated by the collective decisions. Given independence conditions, the utilitarian welfare is maximised by a square root rule, but if voters are risk-neutral and their preferences are sufficiently correlated within constituencies, then a linear rule performs better.

Nicholas Miller's chapter "A Priori Voting Power When One Vote Counts in Two Ways, with Application to Variants of the U.S. Electoral College" investigates the two-tier voting system used to elect the US president from an a priori voting power measurement perspective. He proposes two variants of the Electoral College system and shows how individual voting power may change under these variants. In his "Modified District Plan", a candidate is awarded one electoral vote for each Congressional District and two electoral votes for each State she carries. Under the "National Bonus Plan", a candidate is awarded all the electoral votes of each state she carries (as presently) plus a "National Bonus" of some fixed number of electoral votes if she wins the national popular vote.

Historically, the first authors to propose an analysis of power measurement were Lionel Penrose, who anticipated both John Banzhaf III and the square root rule of two-tier voting systems, and jointly Lloyd Shapley and Martin Shubik. In Part IV, *Penrose, Banzhaf, Shapley–Shubik, et al.*, some of the classical power indices are revisited. The Shapley–Shubik index is a special case of a cooperative game concept, the so-called Shapley value, when applied to simple games—to games where the value of a group of individuals is either 0 or 1. Many alternative indices have been proposed including the Shapley–Owen index, Holler's Public Good Index (PGI) and Schmidtchen and Steunenberg's Strategic Power Index (SPI). In this part, the "inventors" themselves of some of these classic indices provide clarifying comments and justifications.

The classical a priori voting power indices take account of the decision rules together with the vote distribution to measure the influence of voters on the outcomes. However, if an index implicitly assumes that all coalitions are equally likely, it provides a poor indication of the real power distribution. Preference-based indices have been introduced to remedy this defect. In the chapter "Aspects of Power Overlooked by Power Indices", Manfred Holler and Hannu Nurmi argue that, on the contrary, "a priori voting power indices do what they are supposed to do under very special circumstances only" and that the same is true for preference-based indices.

In "Banzhaf–Coleman and Shapley–Shubik Indices in Games with a Coalitional Structure: A Special Case Study", Maria Ekes compares the Shapley–Shubik index and the Banzhaf–Coleman index when a coalitional structure is given. She uses either a composite game structure or games with a priori unions (due to Guillermo Owen) to calculate indices, given a game with 100 voters and various coalition structures.

Holler's PGI has been criticized because it violates local monotonicity (LM). Under PGI a player that makes a larger contribution to winning than another player is not guaranteed to have at least as large as an index value. Holler and Nurmi argue in their chapter “Pathology or Revelation? The Public Good Index” “that cases of non-monotonicity indicate properties of the underlying decision situations which cannot be brought to light by more popular power measures that satisfy LM”. They propose a solution to constrain the set of games representing decision situations such that LM holds for PGI. Considering causality, the authors suggest “that the non-monotonicity can be the result of framing the decision problem in a particular way”. Since PGI was shown to be related formally to the Banzhaf index, can we identify the cause of violation of monotonicity of PGI on the basis of this relation?

In their chapter “On the Possibility of a Preference-Based Power Index: The Strategic Power Index Revisited”, Dieter Schmidtchen and Bernard Steunenberg defend their SPI. They address arguments raised by Braham and Holler regarding preference-based power indices. They tackle the claim made by Napel and Widgren that SPI is not a true power index since it confuses power and luck. The authors close with a reaction to Felsenthal and Machover’s proposition that SPI is a modified Banzhaf index.

In Part V, *Political Competition and Voting Procedures*, two chapters belong to the political competition tradition, but in both cases the authors had power measures in mind. Two chapters are devoted to voting procedures and the last chapter deals with apportionment. (Formally, apportionment is a sort of mathematical dual of proportional representation and can, accordingly, be associated with a voting procedure.)

In the chapter “The Shapley–Owen Value and the Strength of Small Winsets: Predicting Central Tendencies and Degree of Dispersion in the Outcomes of Majority Rule Decision-Making”, Scott Feld, Joseph Godfrey and Bernard Grofman deal with the classical model of spatial voting games with Euclidean preferences. Based on insights derived from the Shapley–Owen value, they explain why the outcomes of experimental committee majority rule are “overwhelmingly located within the uncovered set”. They argue that it is not membership in the uncovered set that matters, but the fact that “alternatives differ in the set of their winsets”. A winset of alternative a is the set of alternatives that are majority-preferred to a. “Alternatives with small winsets are more likely to be proposed, more likely to beat a status quo, and more likely to be accepted as the final outcomes than alternatives with larger winsets”, they claim.

Maria Montero’s chapter “Postulates and Paradoxes of Voting Power in a noocooperative Setting” analyses a leading model of legislative bargaining, due to David Baron and John Ferejohn, in the light of the notion of P-power as defined by Felsenthal and Machover. She demonstrates that the Baron–Ferejohn equilibrium, which was known to violate some minimal adequacy postulates, does not respect other essential properties.

Steven Brams and Marc Kilgour propose in “Satisfaction Approval Voting” a new voting system for multi-winner elections called Satisfaction Approval Voting (SAV). In SAV, voters can approve as many candidates as they like. The winners are *not* those who receive the most votes, “but those who maximize the sum of satisfaction scores of all voters, where a voter’s satisfaction score is the fraction of his or her approved candidates who are elected". The authors show that under SAV (1) all strategies are un-dominated, except the counter-intuitive strategy where a least-preferred candidate is approved of; and that, (2) when applied to party-list systems, SAV apportions seats according to the Jefferson/d'Hondt method with a quota constraint.

In "The Structure of Voters' Preferences Induced by the Dual Culture Condition", William Gehrlein and Souvik Roy revisit pairwise majority voting, the existence of Condorcet's paradox and Condorcet efficiency—viz. the conditional probability that the rule selects a Condorcet winner when one exists. They study "the expected relationship between a classical measure of social homogeneity from the literature and both the probability that Condorcet's paradox will be observed and the Condorcet efficiency of voting rules"—including Borda's rule.

A basic assumption of representative democracy is that representatives represent equal number of persons. When elections are organized on the basis of given entities, one has to define the number of representatives for each entity, with each entity having its proportional share of representation. An immediate difficulty arises from the necessity to have integers; consequently, questions concerning remainders must be solved. A somewhat equivalent problem is faced by the so-called proportional representation voting system. Iain McLean's chapter, "Three Apportionment Problems, with Applications to the United Kingdom" reconsiders this apportionment problem, including a digression on the two-tier power difficulty tackled in Part III of this volume, and describes how these issues have been handled by UK policy makers since 1918.

The collection of papers of this volume have been reviewed by a team of referees in addition to the editors. We wish to thank the referees for their help and for their considered comments. We gratefully acknowledge the excellent advice and assistance of Springer's Martina Bihn, Editorial Director for Business, Economics and Statistics, of her assistant, Ruth Milewski and of Sylvia Schneider, Book Team Production Coordinator. Special thanks are extended also to Karthik Kannan Kumar, Publishing Project Manager at SPi Content Solutions—Spi Global, India, whose diligent handling of the production, most carefully attended to throughout the process, greatly facilitated the successful publication of this volume.

We are delighted as editors, on behalf of all the contributors to the volume, to present this Festschrift to Dan Felsenthal and Moshé Machover in honour of their remarkable achievements in the research of voting power and procedures.

An Interview with Dan Felsenthal and Moshé Machover: Biography, Context and Some Further Thoughts on Voting

Rudolf Fara

1 Introduction

The Dan Felsenthal and Moshé Machover research partnership (F&M hereafter) has been one of the most important influences on the modern development of voting theory. The focus of this well-deserved Festschrift to honour their work has been on voting power and its measurement, the subject of their landmark volume, *The measurement of voting power: theory and practice, problems and paradoxes* (MVP). The focus is entirely apt. The book is a remarkable work that played a major role—possibly *the* major role—in resuscitating the voting power field that until its appearance had critically stalled.

When I proposed an 'interview' in lieu of the stock Festschrift biography, Dan playfully questioned whether I was asking them to write the introduction too! My first thought had been to record an extensive interview similar to my earlier video archives on the work of the philosophers Strawson, Quine and Davidson,[1] but new plans precluded this approach at this time.[2] On reflection, Dan's joke wasn't so far from the objective of this interview. F&M were being asked, in effect, to steer the selection of discussion particularly with regards to matters personal. Consider that on our fair planet blue, Dan and Moshé's extensive works, both joint and individual,

[1]Details of these archives from Philosophy International's *DVD Library of Philosophy* are available at: http://www.lse.ac.uk/pi.

[2]A series of recordings of interviews (including in-depth discussions with Dan and Moshé), workshops and other course supplements will form an integral part of the new *Vote*Democracy course that Dan alludes to in his final response in this interview.

R. Fara (✉)
Centre for Philosophy of Natural and Social Science, London School of Economics and Political Science, London, UK
e-mail: r.fara@lse.ac.uk

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare,
DOI 10.1007/978-3-319-05158-1_2,

are just a mouse-click away. A ‘Felsenthal and Machover’ search on any browser serves up 38,500+ hits in a few milliseconds. There are excellent biographical sketches by Hannu Nurmi, for example, in Keith Dowding’s comprehensive *Encyclopedia of Power (2011)*[3]; Moshé’s Wikipedia entry is informative too, particularly about his political work.

With so much information so easily retrievable, the goals set for this interview were modest: to invite some biographical insights and early background to their research partnership unlikely to be found elsewhere; to elicit their views on some of the basic problems that beset the foundations of social choice; to prompt them to field a few specific questions from some of the major voting theorists; and to bring us up-to-date with their current activities.

Dan and Moshé’s involvement in the new *Vote*Democracy[4] project, adverted to by Dan in his final response below, deserves further mention since it is a pedagogical development strongly influenced by over a decade’s work of the Voting Power & Procedures (VPP) research programme at the London School of Economics to which they have contributed enormously. Both have already participated enthusiastically in laying the foundations for this new educational enterprise. In this regard, I am reminded of discussions with Moshé preliminary to the formation of VPP at LSE in 2000 when we agreed, with particular reference to the prospect of EU enlargement, that our research project should have a conspicuous practical dimension. That dimension featured prominently, early and late, from the importance Dan and Moshé ascribed in their MVP to a multidisciplinary approach to research and its practical application, to the continued emphasis on promoting interactivity between academic theorists and practitioners that has characterized more than a decade of VPP research. *Vote*Democracy might be seen as the dynamic culmination to this theory/practice research approach: to imbed the study of voting, the empirical heart of representative democracy, into mainstream education. I am confident that with their unique contribution to this ‘masterclass’ the Felsenthal-Machover partnership will excel once again.

The interview that follows was conducted by email. Except where a question was directed specifically to one of them, Dan and Moshé were requested to respond jointly or individually as they chose.

RF *Your book,* The Measurement of Voting Power, *was published in 1998 and your first co-authored paper “After two centuries, should Condorcet’s voting procedure be implemented?” appeared in* Behavioral Science *in 1992. What brought you to work on voting theory and how did your collaboration begin?*

DF Early in my academic career I shifted my research interest from investigating governmental policy-making in the US and Israel in the areas of higher education and public health to the fields of bargaining and voting theory. I realized quite soon that a political scientist, like myself, who worked in these fields, would

[3] Available online at: http://knowledge.sagepub.com/view/power/SAGE.xml.

[4] Email vpp@lse.ac.uk for details of the new project.

benefit very much from collaborating with a mathematician because these fields have a significant mathematical content and hence necessitate mathematical skills. It was therefore only natural for me to try, for many years, to entice Moshé—who had been known to me as a family member, a first-rate mathematician, an outstanding teacher and a political activist—to collaborate with me. My efforts were finally successful when, during one of my annual trips to London in 1989 or 1990, I told Moshé about an article that I had recently read—one about the saw-tooth function phenomenon of what was called 'the quorum paradox'.[5] Moshé assisted me in proving that by breaking ties randomly in decision-making assemblies where not all members are present the quorum paradox will be averted. I of course acknowledged Moshé's assistance in the article that I wrote on this subject (entitled "Averting the quorum paradox" and published in *Behavioral Science* in 1991). Soon thereafter I told Moshé about two other articles that I read—one about probabilistic voting and the other about veto-vote—and he immediately had new ideas as to how these two subjects could be extended by us. So the first two joint articles we published in 1992 were on these two subjects. This was the beginning of a wonderful and fruitful collaboration; the rest is history.

MM I can only answer for myself, since our routes to voting theory were quite different. Danny was working on voting theory for many years. For my part, the short answer is that Danny got me into it. My field of research was mathematical logic and related subjects. But in 1986 I got ill with what was later diagnosed as mononucleosis, which brought about an onset of chronic fatigue syndrome. This lasted a few years during which I could do very little work, and felt quite depressed. Then Danny got me out of this. I had known him for many years as a member of my family (he is married to my cousin) and we sometimes discussed his research. One day he presented me with a mathematical problem arising from something he was working on; I think it was about the effect of breaking ties randomly when the voters are evenly divided. It turned out to be a simple problem in finite combinatorics and probability, and I could solve it quite easily. There is nothing a mathematician likes better than solving a problem in a field other than his or her own. So I was very pleased that I was able to help. This got me hooked, and we started to collaborate. It made a very welcome change in my research work, and I owe this productive turn entirely to Danny.

RF *Typically, how do you conduct your joint research and how is it written up? That is, who does what and when and at what point in the process?*

MM Danny is usually the driving force (not to say slave-driver) as he is very industrious (not to say workaholic) whereas I tend to be work-shy until my

---

[5]The paper Dan is referring to is by J.L. Rodgers, J.M. Price and W.A. Nicewander, "Inferring a majority from a sample: the sawtooth function phenomenon", *Behavioral Science* 30 (1985) 127–133. The quorum paradox is about the probability that decisions reached by a voting body that is not fully assembled (a sub-set) will be the same if it were fully assembled may not increase monotonically as the sub-set enlarges. "Sawtooth" is the descriptive geometric representation of the function in a classical two-dimensional diagram.

interest in something is really aroused. So usually it is Danny who proposes a problem or a project, for example, writing our book on voting power. And often he also writes a first draft or at least an outline. Then I get to work on it, edit it and develop the mathematical technicalities and look after the English style. I send this edited version to him, and he amends it and sends it back to me. And so it bounces back and forth like a Ping-Pong ball until it is completed. Danny usually has the last word, as he is much better than me in spotting typos and other lapses.

I should also add that while I do most of the formal and more abstract mathematical presentation, Danny invents most of the tricky examples, especially counter-examples. As a mathematician, I tend to think abstractly and strive at generalization. But Danny thinks in much more concrete terms, and has a knack of finding counter-examples that illustrate some counter-intuitive point. Often I have a hunch that a counter-example can be found by looking in a given direction, but I am unable to actually find it; but he does. He is also much better than me in doing numerical calculations.

DF Moshé's description of the process we underwent in producing our joint work is accurate, and his description of my share is very generous. I would like to add two things to Moshé's description. First, I have worked with other partners during my academic career, but my collaboration with Moshé was the longest and the most fruitful. This was, among other reasons, due to the fact that Moshé is a very patient partner, and hence despite our different work styles, we always managed to settle whatever (few) disagreements we had. Second, because Moshé and I live in two different countries, the communication between us, from the very beginning of our collaboration, was done almost entirely by e-mail messages, occasionally several messages per day. This mode of communication has the advantage that the messages can be kept; so I once proposed to Moshé that perhaps it would be worthwhile—both from the viewpoint of the history of science as well as for reviving our own failing memories—to look closely again at these e-mail messages (thousands!) in order to learn how our ideas about various subjects developed. This project is still to be undertaken.

RF *My next set of questions to you will be about social choice and voting theory in general, followed by questions from several notable theorists in the field. My questions concern more recent developments in voting theory, specifically as Arrow's Impossibility Theorem affects it*[6]*. Of the four criteria (fairness*

[6]In light of Dan Felsenthal's responses that follow, it is important to note that Kenneth Arrow proved his General Possibility Theorem (usually referred to later as Arrow's Impossibility Theorem) in his PhD thesis (1950). In August of the same year it was published in the *Journal of Political Economy* in the paper referred to by Dan. *Social Choice and Individual Values*, Arrow's famous monograph based on and named after his PhD thesis, was published by John Wiley in 1951. In these publications, it must be noted, Arrow proposed five conditions that a fair voting system should satisfy: non-dictatorship, unrestricted domain (or universality—the statement of the condition had a defect in the 1951 version first noted by Julian Blau in 'The existence of social welfare functions', *Econometrica*, 1957), independence of irrelevant alternatives, positive

conditions)—namely, non-dictatorship, Pareto efficiency, unrestricted domain *and* independence of irrelevant alternatives*—that Arrow states cannot all be met simultaneously in converting the ranked preferences of individuals into a social ranking when voters have three or more options (candidates), do you view each criterion as an equally important component of a unified inquiry? Or, could the criteria themselves be ranked in importance?*

MM I will address these questions in one short comment, as this is not really my field. Perhaps Danny, who has done much work on voting procedures, will care to address these questions in greater detail.

I have expressed my views on voting procedures in a short paper "The underlying assumptions of electoral systems", which is included as Chap. 1 in *Electoral Systems: Paradoxes, Assumptions, and Procedures* (2012) edited by Danny and me.

When it comes to electing a representative assembly, I am in any case in favour of using proportional representation. This not only bypasses the dilemma posed by Arrow's Theorem, but is in my opinion politically preferable: I think personalization of politics is not a good thing. It is better to ask voters to vote primarily for this or that programme or platform rather than for this or that person. I say "primarily" because the two are of course not completely separate. In this way the elected assembly is approximately a microcosm of the entire electorate.

As for electing an individual for a position such as president, I think the principle of majority rule is paramount. For this reason I am in favour of a system that Dan and I discussed in one of our first joint papers. It elects a Condorcet winner, if there is one; and where there is no Condorcet winner, it resolves the tie by a weighted lottery in which the weights are allocated according to a uniquely determined optimal distribution.

RF *Interviewing Kenneth Arrow in 1987, the social choice theorist, J S Kelly asked what outstanding social choice problem Arrow would most like to see solved. Arrow responded "reformulating a weakened form of the independence of irrelevant alternatives which stops short of just dropping it completely." By allowing chains of transitivity over irrelevant alternatives, Arrow—rather than expecting as one of Kelly's suggestions, "a deeper impossibility theorem"—concluded that: "I am expecting—no, let me put it more cautiously—I'm* hoping *for a possibility result." Would you care to comment on Arrow here, and on the previous question?*

DF First of all, Arrow in his earliest paper on the subject, 'A difficulty in the concept of social welfare', *Journal of Political Economy*, 58 (4), Aug 1950,

association of social and individual values (or monotonicity) and citizen sovereignty (or non-imposition). In the revised 2nd edition of the book published in 1963, Arrow added a chapter in which inter alia he corrected the point noted by Blau and provided a simpler proof of the theorem with four conditions. A new Pareto efficiency criterion replaced the monotonicity and citizen sovereignty conditions of the original version.

pp. 328–346, has stated five, not four, conditions that a reasonable voting rule which amalgamates the individual preferences into a social choice should satisfy. The fifth condition, which is missing in the above question, was called by Arrow *citizen sovereignty*, which means that the social welfare function should not be imposed.

Of Arrow's five conditions probably the hardest to satisfy is the Independence of Irrelevant Alternatives (IIA) condition. This condition requires that, given the voters' (ordinal) preference orderings among three or more alternatives of which one must be selected, then if a voting rule one employs selects alternative x, it must continue to select x if, *ceteris paribus*, one of the other alternatives is no longer available, that is, becomes irrelevant. Almost all non-dictatorial voting rules do not satisfy this requirement, so it is no wonder that Arrow himself, in his response to J S Kelly quoted above, would have loved to see a weakening of this requirement. In view of the many impossibility theorems formulated so far in social choice theory, I certainly share Arrow's hope for the discovery of a (deep) possibility result.

RF *There are various approaches taken to address the challenges to democratic theory posed by Arrow's Impossibility Theorem, from questioning the premises and methodological assumptions, and attempting to relax one or more conditions, to rejection of the Arrovian approach altogether. Which, if any, of the proposed theory revisions that you are aware of do you think show most promise? Or, should we be looking for a new approach altogether to the social choice problem?*

DF Given that there are no voting rules which can satisfy Arrow's five conditions for electing a single candidate out of three or more candidates, the challenge to democratic theory is to reach a wide consensus among social-choice theorists as to one or more second-best, or tolerable, voting rule(s) for electing a single candidate when the social preference ordering contains a top cycle. I believe that several such voting rules do exist.

RF *William Riker in his highly influential book,* Liberalism against Populism*, argues that in the light of Arrow's Impossibility Theorem, since no voting system can fairly amalgamate individual preferences to capture the "will of the people", the populist or Rousseauean interpretation of democracy as the embodiment of the public will should be rejected. Riker's aim is to establish grounds for a representative liberal democracy in which voting—"election discipline" by his lights—is a "method of controlling officials and no more". Do you accept Riker's view about voting?*

DF Riker's conclusion that no voting system can fairly amalgamate individual preferences to capture the "will of the people" emanates from the long-known fact that if one must select one out of three or more alternatives by means of voting, then all non-dictatorial voting rules may display cyclical majorities. In which case, in Riker's opinion, it is impossible to state what is the will of the majority of the voters. It seems to me that Riker's conclusion is too pessimistic. The fact that with three or more alternatives the social preference ordering may contain a top cycle, does not necessarily imply, in my view, that it is impossible to amalgamate fairly the individual preferences into a reasonable social choice. After all, many proposals have been made since Condorcet's own proposal as to which alternative ought to be selected when the social preference ordering contains a top cycle, and at least some of these (second-best) proposals look to me quite fair.

RF *While no ideal election system has been found in over sixty years since Arrow's discovery, some procedures are shown to be more prone to particular paradoxes than others. If we adopt the strategy of ranking the paradoxes in order of "seriousness", i.e. evaluating some paradoxes as more unacceptable than others, what would justify such an evaluation? Should likelihood of frequency of occurrence take priority over "seriousness" criteria?*

DF As I stated in a recent article, despite the fact that all single-winner voting procedures are vulnerable to several paradoxes (or pathologies), I think that there is a wide consensus among social choice theorists that not all paradoxes are equally undesirable. Although assessing the severity of the various paradoxes is largely a subjective matter, I hold that voting procedures which are vulnerable to paradoxes which I consider as especially intolerable should be disqualified as reasonable voting procedures *regardless of the probability that these paradoxes may actually occur in real-life elections*. These paradoxes are: not electing an absolute winner when one exists, electing a Condorcet loser (or even an *absolute* loser), electing a Pareto-dominated candidate and lack of monotonicity.

On the other hand, I think that the degree of severity that should be assigned to the remaining paradoxes should depend, among other things, on the likelihood of their occurrence under the voting procedures that are vulnerable to them. Thus, for example, a voting procedure which may display a given paradox only when the social preference ordering is cyclical—as is the case for most of the paradoxes afflicting the Condorcet-consistent voting procedures—should be considered more desirable (and the paradoxes it may display more tolerable) than a procedure which can display the same paradox when a Condorcet winner exists. However, in order to be able to state conclusively which of several voting procedures that are susceptible to the same paradox is more likely to display this paradox, one must know what are the necessary and/or sufficient conditions for this paradox to occur under the various voting procedures. Such knowledge is still lacking with respect to most voting procedures and paradoxes. Without such knowledge it is possible—although quite difficult—to assess reasonably the probability of various paradoxes only by examining real-life elections where voters are required to rank-order the candidates. Several such investigations have been conducted, but most of them were limited to small electorates rather than to nation-wide public elections.

RF (A question for MM) *At a press conference held on 16 November 2000 at LSE to launch* Enlargement of the EU and Weighted Voting in the Council of Ministers, *the first publication of the fledgling Voting Power & Procedures research programme that you co-authored with Dan, you were asked by a diplomat (from the Italian embassy if I recall correctly) how you thought that accurately calculating the voting power of EU member states would add anything* *to the process of political deliberation and bargaining. You said that establishing "a level playing field" for the discussions was the main contribution. Would you explain your point, particularly for the benefit of those who appear threatened by what they see as the intrusion of mathematicians and social choice theorists into an area usually dominated by political scientists and lawyers?*

MM We—Danny and I, as well as other researchers—are not naïve. We are well aware that the politicians who negotiate and agree about the decision-making rules of international bodies such as the EU Council of Ministers are largely motivated by party-political interests and what they regard as national interests. But, whatever their motivations, they can only negotiate and act rationally *in furtherance of these interests* if they are properly informed about the objective properties of the rules they are negotiating about. For this they need expert advice, because some important properties of these rules are counter-intuitive. These properties can only be determined by mathematical investigation; common sense can be highly misleading in these matters. Rules adopted by the politicians on several occasions in the past show clearly that they and their political and legal advisers were sadly ignorant of these matters; these rules make no sense even from the perspective of political self-interest; they have properties that the politicians could not possibly have intended.

RF *I'd like now to present a few questions on behalf of some of our contributors to this volume. I have included the questioner's contextual comment where appropriate.*

Steven Brams (New York University):

SB *How critical are the voting-power paradoxes in disqualifying certain power indices and, more generally, advancing a general theory of voting power?*

MM Paradoxes are extremely important in the theory of voting power (as they are in the theory of electoral rules, and indeed in other parts of mathematics, such as set theory). They have advanced the theory both destructively and constructively. What I mean by their "destructive" contribution is that a paradox afflicting a measure or index of voting power may reveal an unacceptable pathology, which disqualifies that measure. But a paradox may also make a constructive contribution. It may turn out that it is not at all pathological, but is rather a counter-intuitive property that reveals something important about the given index and tells us how to use it correctly and not misuse it.

Early in our collaboration, Danny and I discovered that the relative Banzhaf index is afflicted by the bloc and donation paradoxes. The former seemed to imply that in some cases a voter loses voting power as a result of annexing other voters. The latter seemed to imply that in some cases a voter might gain voting power by donating some of his or her voting weight to another voter. At first we thought that this disqualifies the Banzhaf index, but eventually we understood that it is not a pathological property of that index but says something vitally important about it, and the way it had been widely misunderstood and misused.

DF The “destructive” contribution of paradoxes is twofold. First, as Moshé has already stated, revealing an unacceptable pathology from which a measure or an index may suffer, which disqualifies that measure. This is what happened to the Shapley-Shubik (S-S) index after Moshé and I discovered (in collaboration with William Zwicker) that this index suffered from what we called ‘the added blocker paradox’, i.e. that, ceteris paribus, adding a new blocker (vetoer) to every minimal winning coalition may change the ratio between the voting powers of the other (old) voters in the assembly. It seems to me interesting to mention that before discovering this pathology of the S-S index, Moshé and I had regarded the S-S index as the only paradox-free index!

Second, the “destructive” aspect of paradoxes also includes the “destruction” of alleged paradoxes, i.e., showing that they are in fact not genuine paradoxes but reasonable properties of a power index. Thus Moshé and I ‘destroyed’ the ‘paradoxes’ that are known in the literature as ‘the paradox of large size’, ‘the paradox of redistribution’, ‘the paradox of new members’, and ‘the paradox of quarrelling members’.

SB *Can acquiring greater voting power be a curse—in particular, by inducing other players to gang up against you, forming an opposing coalition? Is there evidence that this actually happens?*

DF & MM Voters within an assembly may react to the greater voting powers of other voters in a variety of ways: several voters may decide to form what we called in two of our joint papers ‘feasible or expedient (stable) alliances’; a single (or several) voter(s) may decide to defect from one alliance and join another alliance, or even threaten to withdraw from membership in the assembly if its decision rules were not changed. As far as we know, such reactions seem to have actually happened both in the Council of Ministers of the European Union—where Luxembourg’s consent to have formally no voting power during the first period of the EU can be explained by its joining a feasible alliance with the two other Benelux countries—as well as by the behaviour of various member countries, at different times, in their attempts to reform the decision rules in the International Monetary Fund.

Bernard Grofman (University of California, Irvine):

BG *There are several questions whose understanding Dan and Moshé have greatly contributed to that have been inspirations for my own work.*

Under what circumstances will weights and power scores come into close alignment?

MM This topic is quite well researched, especially with regards to the Penrose-Banzhaf index. By the way, this was one of the two projects that, when Danny and I finished writing our book, I considered as most interesting mathematically and needing to be looked into. (The other project was integrating the structural theory of simple voting games in the edifice of mathematics at large, as viewed by category theory. This has been done quite recently by my former PhD student, Simon Terrington. We have a joint paper about it in the pipeline.)

Penrose more or less assumed (on the basis of rough considerations based on the Central Limit Theorem of probability theory) that in weighted voting with many voters—technically, when the number of voters tends to infinity—and with quota equal to half the total weight, the voters' voting powers tend to be closely proportional to their respective weights. Some simulations that I did with colleagues in Singapore confirmed this tendency. (Chang et al., "L S Penrose's limit theorem: tests by simulation" *Mathematical Social Sciences*, 2006). Innes Lindner provided, at my suggestion, a rigorous proof of this tendency, subject to some natural conditions—the most important of which is that all the relative weights tend to 0. (Lindner and Machover, "L S Penrose's limit theorem: proof of some special cases", *Mathematical Social Sciences*, 2004). And there have been further papers by Lindner and others along these lines.

In fact, approximate proportionality is maintained if the quota is within a certain range, somewhat greater than half the total weight. Słomczyński and Życzkowski showed—as part of their so-called Jagiellonian Compromise (see http://en.wikipedia.org/wiki/Jagiellonian_Compromise)—that the best fit between relative powers and relative weights is obtained when the quota is given by the approximate formula

$$\frac{1}{2}\left[1+\left(\sum w_i{}^2\right)^{\frac{1}{2}}\right],$$

where the w_i are the relative weights (whose sum is always 1). Since we are assuming that the relative weights tend to 0, this optimal quota tends to 1/2 as the number of voters tends to infinity—which confirms what Lionel Penrose assumed back in 1946.

BG *In real world settings, how close will a priori and a posteriori notions of power be to each other?*

MM While the previous question was easy, as the topic is reasonably well researched, this question is, for me at least, quite tricky, because I know of no way of measuring a posteriori voting power that I find convincing. The same applies, by the way, to actual voting power, which is not quite the same as a posteriori power. But my feeling is that an acceptable measure of a posteriori (or actual) power need not be at all close to a priori power, because the former would be heavily influenced by factors that the latter must ignore.

DF The short answer to this question is that the closer is the a posteriori behavior of voters to their (assumed) a priori behavior, the closer will be the voters' a posteriori voting power to their a priori voting power. In other words, the assumptions underlying the measurement of voters' a priori voting power are that voters act independently of one another, that every voter has an equal probability of voting 'yes' or 'no', and hence that all possible coalitions are equi-probable. So the closer it can be shown that these assumptions hold in some real-world decision-making assembly, the closer will be the a posteriori voting power of its members to their a priori voting power. The real practical problem, however, is how to measure the (a posteriori) degree of the voters' (in)dependence. But once this degree is established, empirically or otherwise, one can invoke notions developed in information theory to determine voters' a posteriori voting power. See in this connection the article by Abraham Diskin and Moshé Koppel, 'Voting power: an information theory approach', *Social Choice and Welfare*, (2010), **34**: 105–119, as well as their article in this volume.

D. Marc Kilgour (Wilfrid Laurier University):

MK *I'm more interested in voting systems than in power measures, but it seems to me that there is a parallel. For voting systems, "paradoxes" are desirable properties that at least some systems fail. No voting system we know of survives all the paradoxes we know of, and maybe no system ever will, as we are apparently much better at inventing new paradoxes than new voting systems.*
Do we have to admit that all systems have flaws, and come up with assessments based on the severity and the frequency of flaws?

DF Following Arrow's seminal Impossibility Theorem it is clear that no voting rule for selecting one or more out of three or more candidates can be made paradox-free. Under proportional representation (PR) too, no method for allocating the 'surplus' is paradox-free as long as one requires that the size of the representative assembly will be smaller than the size of the entire electorate, and that every represented faction in the assembly will have an integer number of votes. So the selection of a voting rule for allocating the 'surplus' (under PR), or for electing single winners (or teams), must be based on a subjective assessment regarding the severity and likely frequency of the various paradoxes.
However, what is true with respect to voting procedures is not necessarily true of all systems, i.e., it is not inevitable that all systems must have some flaws. Thus, for example, although all known P-power indices are susceptible to one or more paradoxes, the Penrose absolute I-power measure (and hence also Banzhaf's relative I-power index) are, so far, known to be paradox-free.

MK *Is there any hope of recommending systems that would be good only for specific kinds of elections?*

DF As far as I know, except for recommending that majority rule should be adopted when selecting one out of two alternatives and the number of voters is odd, not much work has been done for recommending the use of some voting rule(s) only for specific kinds of elections. Perhaps the use of approval voting should be limited only for selecting teams (e.g., committee members), and perhaps some variation of veto-vote should be used when the main purpose (in relatively small electorates) is selecting an alternative such that no voter considers the elected alternative to be his least-preferred and no voter will regret the way s/he voted.

Nicholas R. Miller (University of Maryland Baltimore County):

NM *Because of your digging into the history of voting power measurement, the well-known Banzhaf measure is now often referred to as the Penrose-Banzhaf measure. When and how did you first become aware of the work of Lionel Penrose that relates to voting power?*

MM The short answer is that the credit for recalling Lionel S Penrose's pioneering work on voting power is due to Peter Morriss. When Danny and I started working on voting power we—like almost everyone working in this field—were quite unaware of Penrose's work, and misunderstood the relativized form of his measure, known as the Banzhaf index (after John F Banzhaf, who independently reinvented some of Penrose's theory nearly two decades later), as quantifying some kind of payoff. This is reflected in our first paper on voting power, "Postulates and paradoxes of relative voting power—a critical re-appraisal", *Theory and Decision*, 1995.

The widespread misapprehension prevailing at the time is described and discussed in our 2005 paper "Voting power measurement: a story of misreinvention", *Social Choice and Welfare*.

This misapprehension was also reflected in a paper I coauthored with Garret and MacLean, also published in 1995, "Power, power indices, and blocking power: A comment on Johnston", *British Journal of Political Science*. Pete Morriss responded to that paper and put us right. He was among the very few people who did not share the confusion about the Banzhaf index, and one of the even fewer who knew about Penrose's work; he had written about all this in the first edition of his book *Power: A Philosophical Analysis* (Manchester University Press, 1987). I suppose social-choice researchers had ignored his book, because who has time for philosophy?

I was persuaded that Morriss got it right. At about that time (or perhaps s bit later) Dan and I were working with Bill Zwicker on a paper, which eventually got published in January 1998 "The bicameral postulates and indices of a priori relative voting power", *Theory and Decision*. I told them about Morriss' critique, and they were also persuaded by it. In that paper we first made the terminological distinction between I-power and P-power. The rest is (literally) history.

Stefan Napel (University of Bayreuth):

SN *Papers by Maaser, Napel, Widgrén and others have highlighted that the Shapley-Shubik index nicely picks up influence in settings with interval policy spaces and single-peaked voter preferences that are independent and uniformly distributed à la Penrose. Your Comment 3.2.15 in Felsenthal & Machover (1998) indicates that the Banzhaf measure also has an interpretation as a prize in settings where an outsider has a given willingness to pay for the passage of a proposal. Do you believe that a dichotomous characterization of the Banzhaf measure as "the only reasonable a priori measure of I-power" and the Shapley-Shubik index as "at best a measure of P-power" is nonetheless warranted?*

Of course, this by no means excludes the possibility that these measures have other interpretations and applications. So this is a dichotomy only as far as strictly a priori voting power is concerned.

The papers by Maaser et al. do not contradict our assessment, because they do take voters' preferences, as well as the structure of policies to be decided, into account. Although the (special) assumptions they make about preferences and policies are arguably, or apparently, quite weak, they do make a difference. So, the Shapley-Shubik index may indeed measure some form of I-power that is not strictly a priori.

As for the interpretation of the Penrose-Banzhaf measure as a price or bribe paid by an outsider who is interested in buying a vote: this is not really a matter of P-power, but I-power in thin disguise, which makes it look superficially like P-power. What the outsider buys is really the voter's influence over the outcome, and so of course the price is proportional to the amount of influence.

DF & MM We believe that our characterizations of the Penrose-Banzhaf measure as the only reasonable measure of absolute *a priori* I-power, and of the Shapley-Shubik index as "by far the most serious known candidate" as measure of a priori P-power are warranted. We stress "a priori", which means that voters' preferences and relations of interdependence, as well as the issues to be decided, are totally ignored.

SN *In its justification for awarding the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2012 to Lloyd S. Shapley (and Alvin E. Roth), the Prize Committee—perhaps tellingly—barely mentions the Shapley-Shubik index. It emphasizes the Shapley value's application in cost sharing problems and, mostly, other contributions by Lloyd Shapley's work to improving (our understanding of) the world. What are in your view the most important applications where power measures (of the Shapley-Shubik, Penrose-Banzhaf, or any other variety) have actually had an impact on the real world?*

DF & MM Although the 1954 paper by Shapley and Shubik, in which they introduced their index, was for a long time the most cited paper in political science, in our view the most important application where an a-priori power measure *has actually had an impact in the real world* was Banzhaf's index. See Ch. 4 (esp. §4.2) in our 1998 book.

Hannu Nurmi (University of Turku):

HN *Do you think that democratic systems are inherently majoritarian, i.e., necessarily based on majority rule?*

Follow-up: can you envision circumstances under which special precautions ought to be made from protecting the interests of minorities? If you do, what kind of mechanisms would guarantee just outcomes under the circumstances?

DF The term 'majority decision rule' means that in order to implement a proposed resolution that changes the status quo and obliges all voters to abide by it, a *simple majority* must support it, i.e., slightly more than half, of the voters. Both Condorcet's jury theorem and May's simple majority theorem (1952), provide some theoretical (normative) support to the majoritarian decision rule. However, May's theorem (showing that majority rule is the only binary decision rule that is anonymous, neutral, decisive and monotonic) is limited to the choice of one out of two alternatives when the number of voters is odd, because when more than two alternatives exist or when the number of voters is even no alternative may be supported by a simple majority of the voters. Similarly, according to Condorcet's jury theorem the probability that the jury will reach a 'correct' decision increases with an increase in the number of jurors, but this happens only when all jurors vote independently of one another and each juror's probability of voting for the 'correct' decision is larger than 1/2. However, if jurors do not vote independently of one another, and/or every juror's probability of voting for the 'correct' decision is smaller than 1/2, then the optimal number of jurors is 1.

Given that the *theoretical* (or normative) justification for adopting the majority decision rule is limited, it is not clear whether and when so-called democratic systems should adopt this rule. This is so because there exist two alternative principles which seem no less 'democratic', or desirable, than majority rule, but which cannot be satisfied simultaneously with the principle of majority rule.

The first of these alternative principles is that *all voters should have an equal chance that the elected alternative will constitute their top preference*. To implement this principle one would need to select the winning alternative by lot—as was done in ancient Greece for selecting several types of officials. However, implementing this rule in practice may result, admittedly with small probability, that the elected alternative will constitute the top preference of a very small minority of the voters and the bottom preference of a very large majority of the voters.

The second principle is that *the elected alternative will not constitute any voter's least preferred alternative*. In order to implement this principle one would have to give every voter some degree of veto power, which can range between requiring that any proposed resolution can pass only if it is supported by *all* voters (unanimity)—which may lead, in turn, to paralysis and total inability to change the status quo—or to provide every voter (or group of voters) with only a limited veto power, e.g., the ability to veto only one, or a small number of, the competing alternatives.

So my answer to the question whether the principle of majority rule should always be adopted is 'no'. However, which of the above-mentioned three rules (or a mixture thereof) should be adopted depends on the relative importance one assigns to them both in general as well as in particular situations. Thus, for example, in order to achieve a wide consensus, or reduce the danger that majority rule may lead to a 'tyranny of the majority', or make it more difficult to change the status quo in order to protect that what are regarded as especially important rights or issues will not be determined or overturned by a mere majority, it is customary in most constitutions, as well as in some laws, that some articles can be amended only if the amendment is supported by some super-majority, thus implementing a mixture of the majority/minority principles.

However, I have no definite answer to the kind of chicken-and-egg question, namely which rule (or mixture of rules) should be adopted by an assembly when it must decide which rule (or mixture of rules) it should use when it will have to select one out of three or more alternatives. In practice the rule actually adopted in such situations is the simple majority or even plurality rule, not only because 'tyranny of the minority' is considered far more undesirable than 'tyranny of the majority', but also because it is the most prevalent as well as easiest to enforce (if need be by physical force). After all, God is usually on the side of the big(ger) battalions . . .

Friedrich Pukelsheim (University of Augsburg):

FP *Is the difficulty for the results of voting theory research to reach practitioners a deficiency of the practitioners, of the theoretical material, of diverging voices in the academic community, or of something else?*

DF & MM In our opinion the difficulty for the results of voting theory to reach practitioners and/or to be implemented, can be explained by all the factors you have explicitly mentioned, on which we will not elaborate, plus the following factors which may be included under "something else":

1. The public at large in most countries is normally not interested or concerned with the advantages or disadvantages associated with various voting rules—and hence there is usually no political pressure, in most countries, to reform their voting rules—even when it is clear that they suffer from some serious defect.
2. The active public involvement of voting-theory experts is a necessary condition for advocating and implementing any change in current voting rules. Except for jurists (who presume themselves experts in voting theory), and perhaps except for some few real voting theorists (e.g., you and Steven Brams), most voting theorists shy away from public exposure; hence the results, as well as debates, regarding voting-theory research are confined almost exclusively to the professional literature and professional conferences without affecting reality.
3. *Ceteris paribus*, there always exists the possibility that by replacing or reforming a voting procedure or a decision rule, one would obtain a different outcome—i.e., some current winners may become losers and vice-versa. Since the amendment of most voting procedures, in most countries, requires a super-majority in the legislature—and sometimes also a referendum—it is no wonder that only relatively few countries reformed their voting rules during the last century. Thus, for example, the US has not succeeded to date to abolish its Electoral College, although it is clear, as has already happened several times, that a presidential candidate who is supported by an absolute majority of the voters nation-wide, may not be elected.

Maurice Salles (University of Caen):

MS (A question for Dan Felsenthal) *It has been shown that Kemeny's rule is NP-hard (from a computational complexity viewpoint). This means (it is a joke but not only) that "a candidate's mandate might have expired before it was ever recognized" (quote from a paper by Bartholdi, Tovey and Trick). So Dan, do you still consider that Kemeny's rule is one of the best voting procedures, if not the best?*

DF As I stated in Ch. 3 of the volume edited by Moshé and me and published in 2012, both Copeland's and Kemeny's rules seem to me to be the best two voting procedures for electing a single candidate. In my view this is so because both these procedures are Condorcet-consistent and not susceptible to what seem to me to be especially intolerable paradoxes. It is true that in single-winner elections with n competing alternatives, it is much easier to determine the winner(s) according to Copeland's than according to Kemeny's rule. This is so because under Copeland's rule one must conduct up to n(n-1)/2 pairwise comparisons (which is probably not NP-hard), whereas under Kemeny's rule one must inspect up to n! possible social preference orderings (which may become NP-hard even for a moderately large n).

However, because Copeland's rule is likely to result in a tie and is more manipulable than Kemeny's, perhaps the best (hybrid and non-NP-hard) rule for electing a single candidate would be to elect the candidate with the highest Copeland score, and if there exist several such candidates to break the tie among them by using Kemeny's rule.

Dieter Schmidtchen (University of Saarland) and Bernard Steunenberg (Leiden University):

DS & BS *Why do we need a unified approach to the measurement of power based on non-cooperative game theory?*

DF & MM We are not sure that we do need a unified approach, let alone one based on non-cooperative game theory. There is no unique notion of voting power (let alone power in a more general sense . . .). This applies even to the much more special case of a priori voting power. There was a time when most researchers in the subject thought there was a unique notion of a priori voting power, based on cooperative game theory, and measured by the Shapley-Shubik index. Even Banzhaf, who independently reinvented part of Penrose's theory, which has a very different notion of a priori voting power—not based at all on game theory, strictly speaking—vacillated on this issue, as we pointed out in our 2005 paper "Voting power measurement: a story of misreinvention", *Social Choice and Welfare*. The outstanding exceptions were Coleman, who also reinvented part of Penrose's theory but was unaware not only of his work but also of Banzhaf's, and Morriss, who knew about both Penrose and Banzhaf. We followed Coleman and Morriss in insisting on the conceptual distinction and gave it terminological expression. Perhaps even this distinction is not exhaustive, as has been suggested by Laruelle and Valenciano (see their book *Voting and Collective Decision-Making: Bargaining and Power*, Cambridge University Press, 2008).

DS & BS *To what extent does the concept of power, as developed in political science (or game theory, if you wish), depend on not specifying further details of the situation in which political actors make decisions? That is, is there still space for power if we would further develop, and include in the analysis, other important concepts such as institutions, behavioral regularities and psychological characteristics?*

DF & MM In our work we only deal with a priori voting power, which by definition excludes all information other than the decision rule itself. However, we think, or at least *hope*, that there is scope for a theory of *de facto* voting power (which would take into account factual institutional and other real-life factors), as well as for a theory of *a posterio*ri voting power (which would deduce voters' powers from past data on their voting and the outcomes). Such theories would be of very great interest and importance. But these are notoriously difficult tasks.

RF *In light of the fact that VPP is currently contemplating a new pedagogical project, and that you both still play active roles in the programme, it is unlikely that your 'formal' retirement will hamper your opportunities to participate in your field so long as you feel inclined. However, Stefan Napel has asked whether, with the recent (or within a few years) retirement of a whole generation of scholars of voting theory, are you anxious that much of the work in the field will be forgotten again, as seems to have been the case with much of the work done before your 1998 book appeared?*

DF & MM We are not sociologists of science and hence we do not have a well-researched explanation regarding why knowledge that had been accumulated in some scientific field was sometime forgotten and subsequently reinvented, and then forgotten again, and then again reinvented, and so on. All we know is that, unfortunately, voting theory has undergone such a cycle. We believe that one possible explanation as to why one may have to reinvent knowledge that had been accumulated and thereafter neglected for waning interest in the field, is the difficulty, during some periods in history, to easily retrieve the knowledge that had been accumulated once the interest in the field waxes again. So although it is quite possible that, for various reasons (e.g., the retirement of a large group of scholars who did not manage to raise a sufficient number of disciples, or the growth of new fads in various branches of science), interest in voting theory may wane again. We think it is much less likely that knowledge that has been accumulated so far will vanish and again have to be reinvented when interest in the field re-waxes. This is so because vast amounts of knowledge have been put on the internet and hence it can relatively easily be retrieved. However, it seems to us that nowadays reinvention in various fields of science, including voting theory, may occasionally occur not as a result of lack of information but rather as a result of the growing difficulty to scan vast amounts of relevant information dispersed among a growing number of sources (scientific disciplines).

RF *At what stage in life do you think exposure to the* voting *problems of democracy ideally should begin?*

DF & MM We concur with Donald Saari who advocates that exposure to voting problems of democracy should begin in the fourth grade, perhaps even in the second grade, in elementary schools. See his short paper 'A fourth grade experience' downloadable from the internet at http://www.math.uci.edu/~dsaari/fourthgrade.pdf

RF *What academic background do you think best equips someone wishing to do research in voting theory?*

DF & MM Voting theory is a classic inter-disciplinary branch of science because it has various aspects that are rooted in various disciplines. However, although voting theory should have been of significant interest to political scientists, philosophers and lawyers, most researchers in this field have been, and still are, either mathematicians or economists—because of the important mathematical aspects of this theory, as well as because most political scientists, philosophers and lawyers shun mathematics. However, we think that interdisciplinary teams, e.g., the Felsenthal-Machover firm, would achieve the best research in voting theory.

RF *After more than twenty years of research in the field of voting theory, what do you think has been your most important contribution?*

DF & MM Without a doubt, although we produced a sizable amount of scientific articles, it seems to us that our most important contribution in the field of voting theory was our 1998 book *The Measurement of Voting Power: Theory and Practice, Problems and Paradoxes* which is still widely cited by scholars working in the field of voting power.

RF *Whose work in voting theory, including both historical and more recent, do you most admire?*

DF The two books in voting theory which I most admire—and which initially attracted me to do research in this theory—are Duncan Black's 1958 book *The Theory of Committees and Elections* (Cambridge University Press), and Robin Farquharson's 1969 book *Theory of Voting* (Yale University Press).

MM My highest admiration is for the pioneering work of Lionel S Penrose, who single-handedly invented the theory of voting power, and developed it to a considerable extent.

RF *Dan and Moshé, I would like to close this interview with the last words belonging to you in the form of, say, a question each. What question would each of you have liked to have been asked, but were* not*, and what would have been your respective answers?*

DF & MM Both of us agree that the question we would have liked to be asked, but were not, is our views about recommended future research work in the fields of voting power and voting procedures. Following are our answers to this question.

DF As I have already indicated in my responses to several of the questions we were asked, I would have liked to see additional research done both in the field of voting power as well as in the field of voting procedures. Thus, for example, I would welcome progress in discovering the necessary and/or sufficient conditions for the occurrence of some of the paradoxes afflicting at least some voting procedures; I think it is important that coalition theory will be further extended such that it could be used by the voting-by-veto procedure when the number of voters is larger than the number of competing alternatives, i.e., when only groups of voters can veto some competing alternative(s). This implies, in turn, that better models of sophisticated voting by both individuals and groups in situations of complete information are needed. As I also mentioned, a necessary condition for making further progress in the measurement of a posteriori voting power is the development of a reasonable measure of voters' degree of (in)dependence.

As to my own academic plans at age 75, I think I can use more productively whatever limited skills I still have by engaging in disseminating some of the knowledge regarding voting power and procedures that has already been accumulated than in creating new knowledge. Therefore I, together with Moshé, the editors of this volume and some additional colleagues, are now engaged in developing a novel multi-level pedagogical program, which we tentatively call *Vote*Democracy. This, it seems to me, will be my last venture.

MM As you know, for most of my academic career I was engaged in work on mathematical logic and the foundations of mathematics. Foundational issues have continued to occupy me also after I started to collaborate with Danny on social choice. For my part, what attracted me most in our joint work—especially our book—was clarification of foundational issues, the foundations of the theory of voting power. But, as I mentioned in my response to Bernie Grofman, one important foundational task remained largely untackled until recently: integrating the theory of simple voting games in the edifice of mathematics at large. And, as I also mentioned, this topic has now been addressed in Simon Terrington's PhD thesis and in our recent joint paper. This opens the door to interaction between the theory of simple voting games, in particular voting power, and other, apparently unrelated, parts of mathematics, such as combinatorial topology. This looks to me as a very desirable area of new research.

Part II
Foundations of Power Measurement

The Measurement of Voting Power as a Special Case of the Measurement of Political Power

Abraham Diskin and Moshe Koppel

1 Introduction

Felsenthal and Machover have made substantial contributions to the measurement of voting power. It is worth bearing in mind, however, that the notion of political power is actually a quite general one of which voting power is one instantiation. In this brief paper, we consider political power in the general sense and propose a definition. We will show that when applied specifically to voting power, our definition is a generalization of Banzhaf's definition.[1]

Let's begin by considering some remarks on power in its most general sense. Russell (1938, p. 4) notes that: "... the fundamental concept in social science is Power, in the same sense in which Energy is the fundamental concept in physics". Nevertheless, the precise definition of political and social power remains the subject of controversy even today.

Max Weber, a founding father of the study of power in the twentieth century, suggested a probabilistic approach to the measurement: "Power is the probability that one actor within a social relationship will be in a position to carry out his

[1] Many of the technical aspects of this paper have been developed in greater detail in our previous paper (Diskin and Koppel 2010).

A. Diskin (✉)
Department of Political Science, The Hebrew University of Jerusalem, Jerusalem, Israel
e-mail: mshanna@mscc.huji.ac.il

M. Koppel
Department of Computer Science, Bar-Ilan University, 52900 Ramat-Gan, Israel
e-mail: koppel@cs.biu.ac.il

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare, DOI 10.1007/978-3-319-05158-1_3,

own will despite resistance, regardless of the basis on which this probability rests." (Weber 1978, 1921–1922).

Dahl elaborated on Weber's idea. He emphasized the importance of what A does in order "to carry out his own will", and specifically to the extent to which the probability of the desired outcome depends on A's activities: "Suppose I stand on a street corner and say to myself, 'I command all automobile drivers on this street to drive on the right side of the road'; suppose further that all the drivers actually do as I 'command' them to do; still, most people will regard me as mentally ill if I insist that I have enough power over automobile drivers to compel them to use the right side of the road." (Dahl lived in the United States, not the United Kingdom.) Therefore, Dahl concludes, "My intuitive idea of power then, is something like this: A has power over B to the extent that he can get B to do something that B would not otherwise do". (Dahl 1957).

Our emphasis will be less on A's power over another player than on A's power in determining the outcome, but Dahl's probabilistic approach to power is nevertheless apt. Let's explore it a bit more carefully.

2 Political Power and Probability

Dahl proposes the following formal definition of political power. Assume that some agent has a choice between two courses of action, a1 and a2 and that the probability of some desired outcome occurring is p1 if he does action a1 and p2 if he does action a2. Then, his power in determining the desired outcome is $|p1 - p2|$.

This definition has some intuitive appeal. For example, in Dahl's fanciful example, the agent's two courses of action, commanding or not commanding drivers, yield identical probabilities of the desired outcome. Hence, by Dahl's definition, the agent has no power.

Dahl's definition suffers from a number of weaknesses. It does not generalize to the case where there are more than two possible courses of action. It does not distinguish between the cases in which the agent can shift the probability from 0.8 to 1 and the case in which he can shift the probability from 0.4 to 0.6. It relates only to a specific state of the world in which all other influences on the outcome (but the agent's course of action) are determined, but not to the more natural case in which other agents actions are known only up to some probability. Finally, it does not relate to the probability that the agent will choose a particular course of action. Imagine, for example, in a variation on Dahl's fanciful story, the agent's desired outcome is for drivers to drive on the left side of the road and one possible course of action is to take out a sub-machine gun and command all drivers to comply. While Dahl might suggest that the agent has complete power over the outcome, we might prefer to take into account the unlikelihood of this course of action and assign the agent less than full power.

Let's consider a different definition of power based on probabilities of outcomes that is somewhat more sophisticated than that of Dahl.

Our basic idea is that instead of using the *difference between probabilities* as the basis for the measurement of power, we should use the *diminution of uncertainty*.

Let $s_i \in S$ be a possible state of the world and let $p(d|s_i)$ denote the probability of outcome d, given that the world is in state s_i. We are interested in measuring our uncertainty about the outcome d, given that we are in state s_i. For example, if $p(d|s_i)$ is either 0 or 1, then we have no uncertainty at all. Conversely, if $p(d|s_i)$ is 1/2, then our uncertainty is maximal. Before we get into the technical details regarding how this can be done, let's see why this is relevant to our problem.

Suppose that we get an additional item of information about the world. We are told what action is taken by agent a from among the range of possible actions $\{a_i, \ldots, a_m\}$. Imagine that we could precisely measure how much that new information decreases our uncertainty about the outcome d. Our claim is that *this decrease is precisely the power of the agent a*.

To see this, consider the extreme case where the probability of d occurring is 1/2 (that is, maximally uncertain) when a's action is unknown, but becomes either 0 or 1 (that is, not uncertain at all) once we know a's action. Then a has complete power over the outcome d. On the other hand, if we know the outcome with certainty without knowing a's action or, more generally, if a's action does not affect the probability of any outcome, then a has no power at all.

Formally, the measure of uncertainty is well understood from information theory and is captured by the function $f(p) = -p * \log p - (1-p)\log(1-p)$. As can easily be seen, this function is consistent with the extreme cases noted above. Now let $p(s_i)$ denote the probability that the world is in state s_i. Then the uncertainty regarding the outcome d is simply the weighted average of that outcome over all possible states $s_i \in S$. Formally, let D be a random variable that takes the values {d occurs, d does not occur} and let S be a random variable that takes the values $\{s_1, \ldots, s_m\}$. Then the average uncertainty of D given the value of S is denoted by $H(D|S) = \Sigma_i p(s_i) * f(p(d|s_i))$. Let A be a random variable that can take the values $\{a_1, \ldots, a_m\}$ representing the possible actions of agent a. Similarly, the average uncertainty of D given S and A is denoted by $H(D|S, A) = \Sigma_i p(s_i, a_i) * f(p(d|s_i, a_i)$. The power of agent a (Power(a)) is simply $H(D|S)–H(D|S, A)$.

It can easily be seen that this formula yields the intuitive result for each of the extreme cases considered above. By definition, Power(a) lies in the range [0,1]. Power(a) = 0 when the knowledge of a's action does not remove any uncertainty concerning the outcome. Power(a) = 1 when there is total uncertainty concerning the outcome when a's action is unknown, but there is full certainty concerning the outcome when a's action is known.

Note that our definition is trivially generalizable to the case where the possible outcomes are not limited to the values {d occurs, d does not occur}, but rather consists of any finite number of possible outcomes.

3 Voting Power

The problem of voting power, in which we wish to measure the power of voter a, offers a particularly neat instantiation of the more general problem considered above (Miller 1999). In the typical voting scenario, we take advantage of the following simplifying assumptions:

1. The set of possible actions by a voter consists solely of voting for or against some proposition.
2. The states of the world consist solely of specifications of how voters other than a vote.
3. The probability of each state of the world is identical.
4. Once the votes of all voters (including v) are known, the outcome is known with certainty.
5. The probability that v will vote for a proposition is 1/2.

Applying these assumptions to the definition of power given above, we obtain the following:

$$\mathrm{H(D|S) - H(D|S, A)} = \Sigma_i p(s_i) * f(p(d|s_i) - \Sigma_i p(s_i, a_i) * f(p(d|s_i, a_i) \tag{1}$$

$$= \Sigma_i p(s_i) * f(p\,(d|s_i) \tag{2}$$

$$= \Sigma_i (1/2)^m * f(p(d|s_i)) \tag{3}$$

Equation (1) holds by definition, Eq. (2) follows from Assumption 4 and Eq. (3) follows from Assumptions 2 and 3.

Now note that, by Assumptions 4 and 5, $p(d|s_i)$ can take only the values 1/2 or 0 or 1. The first occurs when the outcome depends on the vote of a and the latter occur when the proposition fails or succeeds (respectively), regardless of the vote of a. In the first case, $f(p(d|s_i)) = 1$ and in the other two cases, $f(p(d|s_i)) = 0$. It thus follows that

$$\Sigma_i (1/2)^m * f(p(d|s_i)) = K/2^m \tag{4}$$

where K is the number of cases where the vote of a is decisive.

But this last value is of course simply the Banzhaf measure of voting power! Thus we conclude that our general measure of voting power is a generalization of the Banzhaf measure. This fact is of particular interest because it implies that we can measure voting power even in cases where not all the above assumptions hold. In particular, we will consider cases where Assumptions 3 and 5 do not hold.

4 Generalized Measure of Voting Power

Imagine a parliament of 101 delegates in which three political parties are represented: A with 50 seats, B with 49 seats and C with 2 seats. Each of the parties votes *en bloc* either ‘yes’ or ‘no’ with no abstentions possible and with a majority of 51 votes necessary to either pass or block any resolution. In this case, any pair of parties, regardless of size, can get a bill passed or blocked. In this case, both Banzhaf and our proposed measure assign each party equal power.

Now consider the special case in which A and B might support or oppose a bill but that they never agree. Suppose further that C votes half of the time with A and half of the time with B. Under such conditions, it is apparent that C is the only player to have power: its decision always dictates the outcome.

Thus, although we intuitively know how much power each player in this game should be assigned, the Banzhaf measure (1966, 1968) cannot be applied, since it is applicable only in cases where Assumptions 3 and 5 above hold. Clearly, in this case, these assumptions do not hold. Applying our generalized definition, the power of voter C is

$$\Sigma_i p(s_i) * f(p(d|s_i) - \Sigma_i p(s_i, a_i) * f(p(d|s_i, a_i) = \Sigma_i p(s_i) * f(p(d|s_i) \tag{5}$$

$$= 0 + 0 + \frac{1}{2} * 1 + \frac{1}{2} * 1 \tag{6}$$

where the first two terms in Eq. (6) correspond to the cases where A and B both vote in favor or against the proposition, respectively and the latter two correspond to the cases where only A and only B vote in favor of the proposition, respectively. The first two cases occur with probability 0 according to the terms of our story, while the latter two occur with equal probability (hence the 1/2) and $p(d|s_i) = 1/2$ in each case (since it depends on the vote of C). Thus, the measure of C’s power is 1, exactly as should be the case.

By similar calculations, we obtain that the power of A and B, respectively, are both 0.

In Diskin and Koppel (2010), we show many examples where our generalization of the Banzhaf measure yields the intuitively correct answer for voting scenarios in which Assumptions 3 and 5 above do not hold.

5 A Second Measure of Power

Let’s now consider a case for which our measure is not by itself adequate to capture the distribution of power among voters.

Suppose we have a committee of five equal voters and a majority-wins system, where the probability of any set of exactly three (or exactly two) out of five is nil. Each voter votes ‘yes’ or ‘no’ with no abstentions possible. That is, there are no possible bipartitions for which any single voter is decisive. (Let's call this the "no close calls" case.) The Banzahf measure does not apply (since Assumptions 3 and 5 do not hold), but its trivial generalization would assigns every voter the value 0, since there are no cases in which that voter's vote is decisive. Nevertheless, it seems plain that that the players must have *some* power (Machover 2007). Unfortunately, our measure also assigns 0 power to each player.

Let's briefly analyze why this anomaly occurs and how it can be resolved by distinguishing between two kinds of power. Let's first consider what answer we would prefer for the "no close calls" case. It is not hard to see that *any* answer would be somewhat counterintuitive. To see why, consider the following two scenarios.

- Scenario 1: Voter v is an extremely persuasive politician and therefore always succeeds in persuading at least three of the other four voters of his view.
- Scenario 2: Voter v and whoever is sitting closest to him are both very impressionable and once they know the majority view among the other three voters, they always vote accordingly.

How much power shall we assign to v in each of these cases? Perhaps v should be assigned much power in the first case and little power in the second case? But note that in our problem description above, we are given only the probability of each bipartition and the result in each case; we deliberately ignore the question of the dynamics that create such dependencies. The "no close calls" case can be instantiated by either one of these scenarios. Thus, there could not possibly be a single "right" answer to the question of how much power v has. In fact, there are actually two kinds of power.

The measure of power we provided above identifies power with the answer to the following question:

1. *Once we know how everyone but v has voted*, *how much uncertainty remains regarding the outcome*?
   In Diskin and Koppel (2010), we called this kind of power "control". But as is made evident by the "no close calls" case, we might also consider another closely related question:
2. *How much uncertainty regarding the outcome is removed once we know how* (*only*) *v votes*?
   In Diskin and Koppel (2010), we called this kind of power "informativeness".

Intuitively, informativeness complements our previous definition of power. In the "no close calls" case, once we know how all the others have voted, there is no doubt left as to the outcome. Thus, we would assign each voter 0 power. But, as Machover points out, each voter seems to have some kind of power. Our claim is that what each voter actually has is "informativeness": once we know how any individual voter votes, the probability that the outcome will be in accord with that vote is extremely high.

The formal definition of informativeness in information-theoretic terms is given in Diskin and Koppel (2010) and we do not burden the reader with the details here. For the curious reader, we note only that in the "no close calls" case, the informativeness assigned to each voter is $1 - \mathrm{f}(15/16) \approx 0.66$.

6 Conclusion

The study of political power in general, and voting power in particular, have raised many controversies and ambiguities.

Our approach here is to define political power in a very general sense and then to apply the general definition to the special case of voting power. Following Dahl (1957), we define an agent's power in terms of changes in the probability of an outcome depending on the agent's actions. Specifically, we propose that the measurement of an agent's power should focus on the degree that knowledge of that agent's decisions diminishes uncertainty regarding the outcome.

This general measure yields the intuitively correct answer for a number of cases where such an answer is not controversial. Furthermore, for the special case of voting power, this measure coincides with Banzhaf's measure.

Our measure applies also to cases of voting power in which different bipartitions occur with different probabilities, thus overcoming an objection raised earlier by Albert (2003, 2004) and responded to by Felsenthal and Machover (2005). Retrospectively, it seems that in spite of some criticism, the study of voting power is well-rooted in theoretical studies of political power.

The approach suggested in this article generalizes previous approaches along a number of dimensions: it is relevant to any finite number of "outcomes", "states of the world" and "actions" to be taken by an agent whose power we wish to measure. Indeed, nothing about our definition of power limits it to a social or political context; rather, it can be thought of as a measure of power for any kind of interacting probabilistic agents, even nodes in a neural net (Heckel et al. 2013).

References

Albert, M. (2003). The voting power approach: Measurement without theory. *European Union Politics, 4*(3), 351–366.

Albert, M. (2004). The voting power approach: Unresolved ambiguities. *European Union Politics, 5*(1), 139–146.

Banzhaf, J. F., III. (1966). Multi-member electoral districts – do they violate the one man, on vote principle? *The Yale Law Journal, 75*, 1309–1338.

Banzhaf, J. F., III. (1968). One man 3,312 votes: A mathematical analysis of the electoral college. *Villanova Law Review, 13*, 304–332.

Dahl, R. A. (1957). The concept of power. *Behavioral Science, 2*(3), 201–215.

Diskin, A., & Koppel, M. (2010). Voting power: And information theory approach. *Social Choice and Welfare, 34*, 105–119.

Felsenthal, D. S., & Machover, M. (2005). Voting power measurement: A story of misreinvention. *Social Choice and Welfare, 25*, 485–506.

Heckel, R., Schober, S., & Bossert, M. (2013, May). Harmonic analysis of Boolean networks: Determinative power and perturbations. *EURASIP Journal on Bioinformatics and Systems Biology*.

Machover, M. (2007). Discussion topic: Voting power when voters' independence is not assumed. LSE research online, pp. 1–4. http://eprints.lse.ac.uk/2966/.

Miller, N. (1999). Power in game forms. *Homo Oeconomicus, XVI*(2), 219–243.

Russell, B. (1938). *Power: A new social analysis*. London: George Allen and Unwin.

Weber, M. (1978, 1921–1922). In R. Guenther & C. Wittich (Eds.), *Economy and society: An outline of interpretive sociology*. Berkeley: University of California Press.

On the Measurement of Success and Satisfaction

René van den Brink and Frank Steffen

JEL Classification: C79, D02, D71

1 Introduction

In their volume "The Measurement of Voting Power" Felsenthal and Machover (1998) point out that the idea of the measurement of *success* has been part of the theory of voting power since the first scientific contribution in the field by Penrose (1946) and since then has been rediscovered by several other scholars.

We would like to thank Matthew Braham, Keith Dowding, William Gehrlein, Manfred Holler, Serguei Kaniovski, Dennis Leech, Moshé Machover, Peter Morriss, and Stefan Napel for comments and discussions. Considerable parts of the research contained in this paper were already developed between 2005 and 2006 when Frank Steffen was at Tilburg University under a Marie Curie Intra-European Fellowship within the 6th European Community Framework Programme. He gratefully acknowledges this financial support.

R. van den Brink
Department of Econometrics and Tinbergen Institute, VU University, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands
e-mail: jrbrink@feweb.vu.nl.

F. Steffen (✉)
Faculty of Law, Business and Economics, University of Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany

Department of Economics, Finance and Accounting, The University of Liverpool Management School (ULMS),Chatham Street, Liverpool L69 7ZH, UK
e-mail: frank.steffen@uni-bayreuth.de

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare,
DOI 10.1007/978-3-319-05158-1_4,

Despite this fact the notion of *success* has received relatively little attention for a number of decades as—to make use of Laruelle et al.'s (2006) terminology—it was often regarded to be "just a sort of appendix" of *power*. The most prominent early supporters of this view are said to be Dubey and Shapley (1979) who have proven that there exists a linear relationship between *power* measured by the Banzhaf (1965) measure and *success* measured by their "agreement index", which is an extended version of the "Rae (1969) index" (see, for instance, Laruelle et al. 2006).[1] More contemporary supporters are Hosli and Machover (2004) who accentuate that the notion of *success* is "virtually identical" with the notion of *power* and that both "differ only in using a different scale of measurement". Only recently, Laruelle and Valenciano (2005, 2008) and Laruelle et al. (2006) have emphasized and vindicated the relevance of *success* for the normative assessment of collective decision-making mechanisms and have demonstrated that the aforementioned view is misleading.

While their contribution deserves to be acknowledged, they still continue to contribute to an ongoing common confusion in the voting power literature: they appear to regard *satisfaction* to be a synonym for *success*. A careful re-examination of the voting power literature with respect to both notions leads to the following conclusion. In the literature we can find two outcome related correspondences being applied: one which relates an actor's "action" (or "vote") to the collective outcome, let us call this the "action-outcome correspondence" (AOC), and a second one which replaces "action" by "inclination" (or "preference") and which we will call the "inclination-outcome correspondence" (IOC). Both are used to define what authors claim to be *success* or *satisfaction*, respectively. Moreover, in the literature we can find a third notion being called "(individual-group) agreement" which is also used as a label for both correspondences. Note that in comparison to the other two notions, i.e., *success* and *satisfaction*, this one is "neutral" as it leaves it unspecified whether the agreement on the individual level refers to an action or an inclination.

An overview of the combinations of notions and correspondences applied in the literature by different authors is given in Table 1.[2] In addition to the information included in Table 1 the following should be noted. In this strand of literature the usual core point of reference is Rae (1969). He refers to the IOC by speaking about what an actor "would like to have" with respect to the collective outcome, but does not introduce a certain notion for this correspondence. Given the nature of his analysis, in our view, the most appropriate notion appears to be *individual-group agreement*. This notion is applied by Brams and Lake (1978) when they refer to Rae (1969). However, when referring to Rae (1969) starting with Straffin (1977) most authors in the field are careless as they (partly implicitly) claim (1) that Rae (1969) is referring to the AOC and/or (2) that he introduced the notion

[1] Note that Dubey and Shapley (1979) use the notion of "agreement" instead of *success* and that a proof of this relationship can also be found in Brams and Lake (1978) who apply the notion of *satisfaction*. The relationship between these three notions will be clarified in the course of this paper.

[2] While Table 1 includes the most prominent sources, we do not claim completeness.

Table 1 Success and satisfaction in the voting power literature

Correspondence	Notion	Source
AOC	*Success*	Laruelle and Valenciano (2005, 2008) Laruelle et al. (2006)
	Satisfaction	Brams and Lake (1978) Nevison (1979, 1982) Straffin et al. (1982) Laruelle and Valenciano (2005, 2008) Laruelle et al. (2006)
	(Individual-Group) Agreement	Straffin (1977) Dubey and Shapley (1979) Laruelle and Valenciano (2005, 2008)
IOC	*Success*	Barry (1980) Dowding (1991, 1996) Grabisch and Rusinowska (2010)
	Satisfaction	van den Brink et al. (2011, 2013)
	(Individual-Group) Agreement	Rae (1969) Felsenthal and Machover (1998) Hosli and Machover (2004) Straffin (1978)

of *success* and/or (3) that *success* and/or *satisfaction* are appropriate labels for the IOC. Notable exceptions are, for instance, Holler (1982), Felsenthal and Machover (1998), and Hosli and Machover (2004). Finally, the contributions by Barry (1980) and Dowding (1991, 1996) deserve some comments. First, it has to be noted that both do not make any reference to Rae (1969) or any of the other authors listed in Table 1, and that Dowding's (1991; 1996) unique point of reference in this context is Barry (1980). Secondly, it has to be mentioned that Barry (1980), who applies the IOC and uses *success* as a label for this correspondence, adds the additional requirement that the actor in question, in order to be successful, must have chosen an action ("the actor has tried") before the collective outcome occurred.[3]

Having outlined the mix-up in the literature the obvious question of the contribution of its disentanglement arises, i.e., (1) whether this is of historical interest only, and has no consequence for an analysis, or (2) whether there exists an outstanding problem which could be solved only after this exercise has been completed. The brief answer is that both are true. We show that (1) applies if we restrict ourselves to the canonical set-up in the voting power literature, that

[3] We would like to note that Barry (1980) himself defines *success* as the probability to be successful as defined above. A discussion of this issue can be found in Dowding (1991, p. 65). Moreover, note that Barry's (1980) additional requirement is always fulfilled under the canonical set-up as specified in Sect. 2 as under this set-up (1) abstention is not permissible and (2) all actors have to choose their action simultaneously. The issue of "abstention" and its implications for the relationship between *success* and *satisfaction* are addressed in Sect. 6.

is simultaneous decision making, as under this set-up *success* and *satisfaction* coincide, while (2) is correct if we deviate from this set-up, for instance, by allowing for a sequential decision-making mechanism such as the one introduced in van den Brink and Steffen (2008, 2012). It will turn out that *satisfaction* entails *success* as one component.

The remaining paper is organized as follows. Section 2 recapitulates the canonical set-up of a decision-making mechanism in the voting power literature and adds some further assumptions and definitions required to relax the canonical setup in the course of our analysis. In Sect. 3 we investigate the nature of *success* and *satisfaction* and provide general definitions of corresponding measures which are also applicable for sequential decision-making mechanisms. In Sect. 4 we discuss the relationship between *success* and *satisfaction* which requires us to address also the notions of *power* and *luck*. In Sect. 5 we illustrate our results from Sect. 4 by two examples of a sequential decision-making mechanism. Concluding remarks considering the impact of "abstention" on the relationship between *success* and *satisfaction* are contained in Sect. 6.

2 Preliminaries

A *collective decision-making mechanism* (DMM) Γ consists of a decision rule and a decision-making procedure. A *decision rule* is a function which maps ordered sets of individual actions into outcomes, i.e., it states which ordered set of actions generate which outcome. A *decision-making procedure* provides the course of actions of the actors for a collective decision and determines the actions to be counted, i.e., which actions go into the domain of the decision rule.[4]

Assumption 2.1. *For the purpose of this paper let us make the following assumptions regarding* Γ*:*

1. *Proposals submitted to the decision-making body are exogenous: it is the task of the decision-making body either to accept or to reject a proposal, i.e., we have a binary outcome set* $O = \{approval, rejection\}$*.*
2. *A proposal can be submitted to the decision-making body only once.*
3. *The decision-making body contains a finite set of actors:* $N = \{1, \ldots, n\}$ *with* $n > 1$*, whose actions bring about a collective outcome of the decision-making body.*
4. *Each actor* $i \in N$ *has a binary action set:* $A_i = \{yes, no\}$*, where the choice of the* yes*-action means that* i *supports the proposal and the choice of the* no*-action that* i *rejects it.*[5]

[4]For further details, see van den Brink and Steffen (2008, 2012).

[5]This excludes the option of abstention as a *tertium quid*, which can have a considerable impact on the *power* distribution among the actors (see, for instance, Felsenthal and Machover 1997, 1998,

5. *All actors choose their action simultaneously.*[6]
6. *Each actor $i \in N$ has a binary inclination set: $K_i = \{approve, reject\}$ containing i's feasible attitudes towards a proposal.*

Remark 2.2. Note that we treat *inclinations* and *preferences* as synonyms. They express what actors "want" or "like". Hence, they have to be distinguished from the actors' choice of action.

In the voting power literature the canonical set-up of an DMM Γ is based on the first five assumptions of Assumption 2.1. Such a canonical set-up is, usually, represented by a *simple voting game* (SVG), being a pair $(N, \mathcal{W})$ where $\mathcal{W}$ is a collection of subsets called "*coalitions*", which satisfies the following three conditions: (1) $\emptyset \notin \mathcal{W}$, (2) $N \in \mathcal{W}$, and (3) if $T \in \mathcal{W}$ and $T \subseteq T'$, then $T' \in \mathcal{W}$ (monotonicity). A coalition $T \subseteq N$ is said to be *winning* or *losing* according to whether $T \in \mathcal{W}$ or $T \notin \mathcal{W}$. This definition implies that an SVG can also be represented by $\mathcal{W}$ only. Moreover, a coalition T can be regarded as an "index" of the actions of actors who have chosen the same action, for instance, "*yes*" if $T \in \mathcal{W}$. Whether a coalition is winning or losing is determined by the decision rule being applied.

The analysis of the canonical set-up represented by an SVG is primarily based on the membership of actors in coalitions. For this reason van den Brink and Steffen (2008, 2012) have called this approach the "*membership-based* approach". They also introduced an "*action-based* approach" which allows other DMM's, in particular it allows sequential decision making where it can happen that not all actors will be allowed to choose an action. This has already been taken into account for the formulation of Assumption 2.1.[7] Next, we define action and inclination profiles.

Definition 2.3. Given a subset of actors $S \subseteq N$, an *action profile* $a = (a_i)_{i \in S}$ on S is a non-empty ordered set of individually chosen actions $a_i \in A_i$, $i \in S$.

Note that it is not required that an action profile contains an action for every actor. However, all actors have an inclination as expressed in the following definition:

Definition 2.4. An *inclination profile* $k = (k_i)_{i \in N}$ is a ordered set of inclinations $k_i \in K_i$, $i \in N$.

So, an inclination profile k contains one, and only one, inclination k_i for each $i \in N$. We denote the collection of all action profiles on any $S \subseteq N$ by $\mathcal{A}^N$, and the collection of all action profiles containing an action of actor i (i.e. the action profiles on $\{S \subseteq N \mid i \in S\}$) by $\mathcal{A}_i^N$. We denote the actors who are choosing an action in an

2001, and Braham and Steffen 2002). In Sect. 6 we will briefly address the possibility of abstention in the context of the analysis of the present paper.

[6]Note that this implies that a secret but sequential decision-making procedure is permissible as well.

[7]For a detailed description of the "*action-based* approach" we refer to van den Brink and Steffen (2008).

action profile a by $N(a)$, i.e. $N(a) = S$ for $a = (a_i)_{i \in S}$. Furthermore, we denote the collection of all inclination profiles on N by $\mathcal{K}^N$.

Remark 2.5. Assumptions 1, 4 and 5 listed under Assumption 2.1 together with the assumption that agents choose a dominant strategy whenever it exists, imply that under the canonical set-up all actors will always act sincerely, i.e., they will choose an action which corresponds to their own inclinations as it is a dominant strategy for each actor to choose its own first choice (see, for instance, Laruelle and Valenciano 2008, p. 55). If, for instance, an actor i has the inclination to "approve" a proposal it will always choose the *yes*-action. From this we can infer that under the canonical set-up there is no need for a separate modeling of the actors' inclinations, assuming that an actor chooses its action according to its inclination. Hence, its representation by a bare SVG is sufficient.

Remark 2.6. Assumptions 4 and 5 of Assumption 2.1 together imply that all actors are not only allowed, but are also obliged to choose an action which goes into the domain of the decision rule.

Remark 2.6 allows us to specify Definition 2.3 for the canonical set-up as follows:

Definition 2.7. A *simultaneous action profile* $a = (a_i)_{i \in N}$, is a non-empty ordered set of individually chosen actions $a_i \in A_i$, $i \in N$.

So, a simultaneous action profile a contains one, and only one, action a_i for each $i \in N$. What is left to be formulated at this stage are expressions denoting the probabilities that a permissible action profile a and a feasible inclination profile k being compatible with a occurs. The latter is required as for a sequential DMM Γ it can happen, that more than one inclination profile k is compatible with a single action profile a (see Sect. 5).

Definition 2.8. Consider an DMM Γ. An inclination profile k *is compatible with* action profile a if, and only if, for all $i \in N(a)$ it holds that (1) $k_i =$ "approve" if $a_i =$ "yes", and (2) $k_i =$ "reject" if $a_i =$ "no". We denote by $p(a, \Gamma)$ the probability of the occurrence of an action profile a, and by $\tilde{p}(k, a, \Gamma)$ the probability of an inclination profile k which is compatible with a.

We denote the set of all inclination vectors that are compatible with action profile a by $K(a)$.[8] We introduce the notation $k_i \sim a_i$ meaning that the inclination k_i corresponds to the action a_i of actor i in the sense that "$k_i =$ *approve*" if $a_i =$ "yes", and "$k_i =$ *reject*" if $a_i =$ "no". Similar, we use the notation $a_i \sim o$, $o \in O$, meaning that the action a_i corresponds to the outcome o in the sense that $a_i =$ "yes" if "$o =$ *approval*", and $a_i =$ "no" if "$o =$ *rejection*". Finally, we use the notation

[8] So, $p(\cdot, \Gamma)$ is a probability distribution over $\mathcal{A}^N$, while $\tilde{p}(\cdot, a, \Gamma)$ is a conditional probability distribution over $K(a)$.

$k_i \sim o, o \in O$, meaning that the inclination k_i corresponds to the outcome o in the sense that k_i = "approve" if "$o = \mathit{approval}$", and k_i = "reject" if "$o = \mathit{rejection}$".[9]

Remark 2.9. Note that from Remarks 2.5 and 2.6 it follows immediately that under the canonical set-up there exists one, and only one, inclination profile k which corresponds to an action profile a, i.e., $\forall a : \exists \tilde{k} : \tilde{p}(\tilde{k}, a, \Gamma) = 1$, while $\forall k \neq \tilde{k} : \tilde{p}(k, a, \Gamma) = 0$, since $N(a) = N$ in the canonical set-up.

3 Success and Satisfaction

At the outset of this section let us start with a definition of the notions of *success* and *satisfaction*. For this purpose we refer to "Collins English Dictionary" (CollinsDictionary 2009) and "The American Heritage Dictionary of the English Language" (Pinker et al. 2011):

Definition 3.1. *Success* is "the favourable outcome of something attempted".

Remark 3.2. The reference to an "attempt" in Definition 3.1 implies that *success* inherently requires an action. However, it is "inclination-free" as no reference is made to whether that what was attempted was something the actor desired or not.

Definition 3.3. *Satisfaction* is "the fulfilment of a desire".

Remark 3.4. The reference to a "desire" in Definition 3.3 implies that *satisfaction* inherently requires an inclination. However, it does not necessarily require an action as it is feasible that a desire can be fulfilled without having made any attempt at all.[10]

Thus, in general, *success* and *satisfaction* are distinct concepts. Making use of Definitions 3.1 and 3.3 we can now define how *success* and *satisfaction* ought to be ascribed to an actor being a member of a decision-making body under an DMM Γ, which is characterized by the assumptions as listed under Assumption 2.1. We *ascribe success* to i if, and only if, i's chosen action corresponds to the collective outcome.

---

[9]Note that it would be possible to "code" the actions, inclinations and outcomes by 0 and 1, in which case the correspondences $\sim$ defined above could be simply written as equalities. We chose not to do that in this paper to make clear the distinction between actions and inclinations, which is essential for the difference between success and satisfaction (and luck) later on.

[10]This can be illustrated by the following example which we owe to Matthew Braham. Assume that you have the desire to become rich, but you are not doing anything to achieve this. Instead, you are lying on the beach enjoying the sun and the fresh air. However, suddenly one of the seagulls circling over your head drops a valuable diamond which just falls into your lap. Thus, your desire has been fulfilled even you have not attempted anything to achieve this, if we assume that you were not aware of the fact that it might happen that a seagull drops a valuable diamond during the time you are lying on that beach.

Definition 3.5. Let $S \subseteq N$ and $a = (a_i)_{i \in S}$ be an action profile with $a_i \in A_i$ for all $i \in S$. Then, the *success ascription* to $i \in S$, given the collective outcome $o(a, \Gamma) \in O$ is given by

$$\overline{SUC}_i(a,\Gamma) = \begin{cases} 1 & \text{if } a_i \sim o(a,\Gamma) \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Remark 3.6. If the canonical set-up is represented by an SVG, *success* is ascribed to an actor i if, and only if, i is a member of a winning coalition, i.e., if, and only if, $(i \in T \in \mathcal{W})$ *or* $(i \notin T \notin \mathcal{W})$ (see, for instance, Laruelle and Valenciano 2005, 2008).

We *ascribe satisfaction* to $i \in N$, if and only if, i's inclination $k_i \in k$ corresponds to the collective outcome $o \in O$, where the collective outcome results out of an action profile a, which is compatible with the inclination profile k. Note that the collective outcome o is the same for all inclination profiles k which are compatible with the action profile a.

Definition 3.7. Let $k = (k_i)_{i \in N}$ be an inclination profile. Then, the *satisfaction ascription* to $i \in N$, is given by

$$\overline{SAT}_i(k,\Gamma) = \begin{cases} 1 & \text{if } k_i \sim o(a,\Gamma) \text{ for } a \in \mathcal{A}^N \text{ with } k \in K(a) \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Based on Definitions 3.5 and 3.7 we can now formulate corresponding *success* and *satisfaction* measures. The *success* of an actor $i \in N$ in a decision-making body applying an DMM Γ is measured by i's probability to be successful, i.e., by summing-up over i's weighted success ascriptions $\overline{SUC}_i(a,\Gamma)$ for all feasible action profiles $a \in \mathcal{A}_i^N$ where i chooses an action.

Definition 3.8. The success measure *SUC* of a decision-making body applying an DMM Γ is given by

$$SUC_i(\Gamma) = \sum_{a \in \mathcal{A}_i^N} p(a,\Gamma)\, \overline{SUC}_i(a,\Gamma) \;\; \text{for each } i \in N. \tag{3}$$

The *satisfaction* of an actor $i \in N$ in a decision-making body applying an DMM Γ is measured by i's probability to be satisfied, i.e., by summing-up over i's weighted satisfaction ascriptions $\overline{SAT}_i(k,\Gamma)$ for all feasible inclination profiles k.[11]

Definition 3.9. The satisfaction measure *SAT* of a decision-making body applying an DMM Γ is given by

[11]Note that van den Brink et al. (2011) are making use of a version of this measure. However, referring to Straffin et al. (1982) they were unaware of the fact that the correspondence used in Straffin et al. (1982) does not coincide with Rae's (1969).

$$SAT_i(\Gamma) = \sum_{a \in \mathcal{A}^N} p(a, \Gamma) \sum_{k \in K(a)} \tilde{p}(k, a, \Gamma)\ \overline{SAT}_i(k, \Gamma) \ \text{ for each } i \in N. \qquad (4)$$

Remark 3.10. Based on Definitions 3.5–3.9 in conjunction with Remarks 2.5 and 2.6 it is straightforward to see that under the canonical set-up being applied in the voting power literature the values of both measures will always coincide. Hence, there exists no good reason for the usage of two separate measures. As the canonical set-up is, usually, represented by an SVG which relates actions to collective outcomes, but does not include any inclinations, it appears to be reasonable that in this context studies focus on the *success* measure and ignore the *satisfaction* measure. This is, in fact, what we can observe in the recent voting power literature (see, for instance, Laruelle and Valenciano 2005, 2008).

4 Satisfaction, Success, Power, and Luck

After having discussed the notions of *success* and *satisfaction* and their measurement in Sect. 3 their relationship still remains to be clarified. For the purpose of this exercise let us begin with a brief discussion of Barry's (1980) well-known equation "*success* $=$ *power* $+$ *luck*".[12] In Sect. 1 we already addressed Barry's (1980) definition of *success* which according to our analysis in Sect. 3 is, in fact, a definition of *satisfaction*. However, when referring to Barry (1980) in the more recent literature authors have just replaced the IOC, which Barry (1980) originally presupposes, by the AOC (see, for instance, Laruelle and Valenciano 2005, 2008; Laruelle et al. 2006, and, referring to Laruelle and Valenciano 2005, also Rusinowska and de Swart 2006 and Grabisch and Rusinowska 2010). In a similar fashion, when referring to *luck* the same authors have also replaced Barry's (1980) original definition of "luck" by an essentially different one. We will come back to this issue in the course of this section. However, before we would like to specify the notion and measurement of *power* which we apply for our analysis.

The notion of *power* in this paper is based on Braham (2008) and Morriss (1987/2002). In a social context they define *power* as an ability (or capacity) to effect (i.e., to "force" or "determine") outcomes and regard *power* to be a dispositional concept, i.e., *power* exists whether it is exercised or not. Following Braham (2008) we say that an actor i has *power* with respect to a certain outcome if i has an action (or sequence of actions) such that its performance under the stated or implied conditions will result in that outcome despite of the actual or possible resistance of

[12]Note that Barry (1980) does not use the notion of "power" in this context, but refers to *decisiveness*. However, by this notion he means what is usually called "power" in the voting power literature. For a discussion of this issue see, for instance, Dowding (1991, pp. 63–68, 1996, pp. 52–54) or Felsenthal and Machover (1998, p. 41).

at least some other actor. That is, *power* is a claim about what i is able to do against some resistance of others irrespective of its actual occurrence.

We *ascribe power* to an actor in an action profile if, and only if, this actor acts in this profile and by choosing a different action from its action set is able to alter the collective outcome against some resistance of others (represented by those chosen actions of the other actors which are not in line with the "new" action of the actor in question). In this case we say that the actor has a *swing*. With respect to our analysis it needs to be noted that for an DMM Γ the notion of a *swing* as used in the canonical setup is no longer sufficient. This is due to the fact that an actor changing its action in a certain action profile might result in a situation which allows other actors, who did not have an action in the original action profile, to choose an action. On the other hand, it might be that actors who did act in the original action profile cannot act anymore. As a result, it might be that by changing its action an actor might (but not necessarily does) change the outcome. Therefore, we have to distinguish between *strong* and *weak swings* (see Sect. 6, but also van den Brink and Steffen 2008 and 2012 for the case of a sequential DMM).

Let $S \subseteq N$, and $a = (a_j)_{j \in S}$ be an action profile containing action $a_j \in A_j$, $j \in S$. Then, we say that i has a *strong swing* if by altering its choice of action, i forces, ceteris paribus, a new collective outcome.[13] Given action profile $a = (a_j)_{j \in S}$, actor $i \in S$, and alternative action $\hat{a}_i \in A_i \setminus \{a_i\}$, we denote by $\mathcal{A}^a_{i,\hat{a}} = \{\tilde{a} = (\tilde{a}_j)_{j \in T} \in \mathcal{A}^N \mid T \subseteq N,\ i \in T,\ \tilde{a}_i = \hat{a}_i$ and $\tilde{a}_j = a_j$ for all $j \in S \cap T\}$ the set of all action profiles that are possible after actor i changes its action from a_i to $\hat{a}_i$.

[13] As van den Brink and Steffen (2008) point out it is important to draw attention to the interpretation of the ceteris paribus condition in this context. Its common interpretation is that the actions of all other actors remain constant. That is, if i alters its action the only effect that can result out of this is a change in the collective outcome (then we say that i has a *swing* and we ascribe *power* to i). While this "all other things being equal" interpretation is appropriate for simultaneous DMMs, it no longer applies for our more general case of a sequential DMM, which may allow certain actors to exclude other actors from the decision-making as a result of their choices. If we have an action profile and we alter i's choice of action it can happen that the decision-making process requires either the exclusion of actions of other actors from the domain of the decision rule and, hence, from the action profile, or the inclusion of actions by other actors in the domain of the decision rule and, therefore, in the action profile. If such information would be ignored, we can end up with an inappropriate power ascription. In order to avoid this problem we have to go back to the idea behind the literal "all other things being equal" interpretation of the ceteris paribus clause. The basic idea of the ceteris paribus clause is a comparison between two possible worlds: the world as it is (our initial action profile and its associated collective outcome) and the world as it would be if an action were changed (the resulting action profile and its associated collective outcome if i's choice of action were altered). In contrast to the standard interpretation of the ceteris paribus clause our analysis does not necessarily require that all other components of the action profile remain constant after we altered i's choice of action; it requires that the action profiles after the initial change by one actor are consistent with the DMM. This interpretation of the ceteris paribus clause is underlying Definitions 4.1 and 4.2.

Definition 4.1. Let $S \subseteq N$, and $a = (a_j)_{j \in S}$ be an action profile containing actions $a_j \in A_j$, $j \in S$. Then, we say that $i \in S$ has a strong swing in S, if there is an $\hat{a}_i \in A_i \setminus \{a_i\}$ such that $o(\tilde{a}, \Gamma) \neq o(a, \Gamma)$ for all $\tilde{a} \in \mathcal{A}^a_{i,\hat{a}}$.

Let $S \subseteq N$ and $a = (a_j)_{j \in S}$ be an action profile containing an action $a_j \in A_j$, $j \in S$. Then, we say that i has a *weak swing*, if by altering its choice of action, it is, ceteris paribus, feasible that a new collective outcome emerges, but that the outcome does not change necessarily.

Definition 4.2. Let $S \subseteq N$, and $a = (a_j)_{j \in S}$ be an action profile containing actions $a_j \in A_j$, $j \in S$. Then, we say that $i \in S$ has a weak swing in S if there exists an $\hat{a} \in A_i \setminus \{a_i\}$ and $\tilde{a} \in \mathcal{A}^a_{i,\hat{a}}$ such that $o(\tilde{a}, \Gamma) \neq o(a, \Gamma)$, and also there exists an $\hat{a} \in A_i \setminus \{a_i\}$ and $\tilde{a} \in \mathcal{A}^a_{i,\hat{a}}$ such that $o(\tilde{a}, \Gamma) = o(a, \Gamma)$.

Remark 4.3. Note that under the canonical setup of a simultaneous DMM there exist no weak swings. Hence, all swings are strong. For a given action profile $a = (a_j)_{j \in N}$ actor $i \in N$ has a swing in a if $o(\tilde{a}, \Gamma) \neq o(a, \Gamma)$ with $\tilde{a} = (\tilde{a}_j)_{j \in N}$ such that $\tilde{a}_j = a_j$ for all $j \in N \setminus \{i\}$.

Based on Definitions 4.1 and 4.2 we can now extend the usual definition of a swing in order to include the distinction between *weak* and *strong swings*. We do this by fully counting all strong swings and counting weak swings only for a fraction $\epsilon \in [0, 1]$.[14]

Definition 4.4. Let $S \subseteq N$ and $a = (a_i)_{i \in S}$ be an action profile. For $\epsilon \in [0, 1]$, the *power ascription* to $i \in S$, is given by

$$\overline{POW}^{\epsilon}_i(a, \Gamma) = \begin{cases} 1 & \text{if } i \text{ has a strong swing in } a \\ \epsilon & \text{if } i \text{ has a weak swing in } a \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

The *power* of an actor $i \in N$ in a decision-making body applying an DMM Γ is measured by i's probability to have a swing, i.e., by summing-up over i's weighted power ascriptions $\overline{POW}^{\epsilon}_i(a, \Gamma)$ for all feasible action profiles a.

Definition 4.5. For $\epsilon \in [0, 1]$, the power measure POW^{ϵ} of a decision-making body applying an DMM Γ is given by

$$POW^{\epsilon}_i(\Gamma) = \sum_{a_i \in \mathcal{A}^N_i} p(a, \Gamma)\, \overline{POW}^{\epsilon}_i(a, \Gamma) \quad \text{for each } i \in N. \quad (6)$$

[14]Note that van den Brink and Steffen (2008) demonstrate that it is not necessary to specify the value of ϵ for binary DMMs.

Now, in order to link the notion (and measurement) of *power* with the notions (and measurement) of *success* and *satisfaction* as defined in Sect. 3 we will make use of two different notions of *luck*.[15]

The first notion of *luck* is based on Barry (1980) and Dworkin (1981). Barry (1980) has introduced the notion *luck* in the context of studying *power* and defines it as (the probability) of "getting what one wants even when one does nothing", i.e., if "one does not act", where according to our understanding the word "even" is just a careless rhetorical flourish.[16] Barry's (1980) definition of *luck*, we will use the notion *Barry luck* to refer to this, is in line with Dworkin's (1981) notion of *brute good luck* which he distinguishes from good *option luck*. The latter type of *luck* is related to "deliberate and calculated gambles …—whether someone gains or loses through accepting an isolated risk he or she should have anticipated", while *brute luck* is regarded to be "a matter of how risks fall out that are not in that sense deliberate gambles."[17]

Let $S \subseteq N \setminus \{i\}$, $a = (a_j)_{j \in S}$ be an action profile not containing an action from actor $i \in N \setminus S$, and $k \in K(a)$ be a corresponding inclination profile containing inclination $k_i \in K_i$ for actor i. Then, we ascribe *brute good luck* (or *Barry luck*) to i if, and only if, i's inclination k_i corresponds to the collective outcome $o(a, \Gamma) \in O$, but i did not act.

Definition 4.6. Let $S \subseteq N \setminus \{i\}$, $a = (a_j)_{j \in S}$ be an action profile not containing an action from actor $i \in N \setminus S$, and $k \in K(a)$ be a corresponding inclination profile containing inclination $k_i \in K_i$ for actor i. Then, the *brute good luck ascription* to i, is given by

---

[15]Note that, as pointed out by Dowding (1991, p. 64) with respect to one of these notions, both notions must be carefully distinguished from what he call's *personal identity luck*, i.e., "the luck of being the particular person one happens to be" which is discussed by egalitarians (see, for instance, Roemer 1986 or Cohen 1989).

[16]For a critical discussion of the probability requirement in Barry's (1980) *luck* definition see Dowding (1991, p. 65; 1996, 52f). Moreover, we would like to point out an inconsistency in Barry's (1980) analysis. His definition of *decisiveness* being: "the difference between his success [making use of the IOC] and his luck. … it represents the difference that it makes to his success if he tries." Hence, whenever an actor tries according to this definition it is *decisive*, i.e., if i is a dummy actor and i chooses an action, i.e., i tries to get what it wants, i would be decisive. However, this is not what is usually meant by the notion of *decisiveness* (see Footnote 11) and Barry (1980) himself later in his essay writes that *decisiveness* means to be "critical", i.e., to have a *swing*. Now one might argue, that Barry (1980) has meant this and that a dummy actor by definition cannot "try", but this would mean that the notion of a "try" presupposes the ability to be successful with the "try". However, this contradicts also the very basic meaning of a "try" as being just an attempt in order to achieve something, whether one has the ability to do so or not. Taking this criticism into account one could re-define Barry's (1980) definition of *luck* to be: "getting what one wants if one does not try or if one tries without being critical with respect to the action profile in question". Note that the second part of this definition is the definition of *action luck* as contained in Definition 4.8.

[17]Dworkin (1981) illustrates the difference between both types of *luck* by an example of *bad luck*: "If someone develops cancer in course of a normal life, and there is no particular decision to which we can point as a gamble risking the disease, then we will say that he has suffered brute bad luck. But if he smoked cigarettes heavily then we may prefer to say that he took an unsuccessful gamble", i.e., he has suffered *bad option luck*.

$$\overline{BGL}_i(k, a, \Gamma) = \begin{cases} 1 & \text{if } k_i \sim o(a, \Gamma) \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

Remark 4.7. Remark 2.6 implies that under the canonical set-up *brute good luck* does not exist as all actors are obliged to choose an action.

The second notion of *luck* which we call *action luck* is based on Laruelle and Valenciano (2005, 2008). Let $S \subseteq N$ and $a = (a_j)_{j \in S}$ be an action profile containing actions $a_j \in A_j$, $j \in S$. Then, we ascribe *action luck* ($\overline{AL}$) to $i \in S$ if, and only if, (1) i acts in action profile a and (2) i's chosen action a_i corresponds to the collective outcome $o(a, \Gamma) \in O$, but i has no swing in a.

Definition 4.8. Let $S \subseteq N$ and $a = (a_j)_{j \in S}$ be an action profile containing actions $a_j \in A_j$, $j \in S$. Then, the *action luck ascription* to $i \in S$, is given by

$$\overline{AL}_i(a, \Gamma) = \begin{cases} 1 & \text{if } a_i \sim o(a, \Gamma), \text{but } i \text{ has no swing in } a \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

Remark 4.9. Note that Laruelle and Valenciano (2005, 2008) mistakenly claim that their notion of *luck* (which we called *action luck*) is the notion of *luck* which has been proposed by Barry (1980).

Based on Definitions 4.6 and 4.8 we can now formulate corresponding *brute good luck* and *action luck* measures. The *brute good luck* of an actor $i \in N$ in a decision-making body applying an DMM Γ is measured by i's probability to have brute good luck, i.e., by summing-up over i's weighted brute good luck ascriptions $\overline{BGL}_i(k, a, \Gamma)$ for all feasible action profiles a where i does not act.

Definition 4.10. The brute good luck measure *BGL* of a decision-making body applying an DMM Γ is given by

$$BGL_i(\Gamma) = \sum_{a \in \mathcal{A}^N \setminus \mathcal{A}_i^N} p(a, \Gamma) \sum_{k \in K(a)} \tilde{p}(k, a, \Gamma) \overline{BGL}_i(k, a, \Gamma) \quad \text{for each } i \in N. \tag{9}$$

The *action luck* of an actor $i \in N$ in a decision-making body applying an DMM Γ is measured by i's probability to have action luck, i.e., by summing-up over i's weighted action luck ascriptions $\overline{AL}_i(a, \Gamma)$ for all feasible action profiles a where i does act.

Definition 4.11. The action luck measure *AL* of a decision-making body applying an DMM Γ is given by

$$AL_i(\Gamma) = \sum_{a \in \mathcal{A}_i^N} p(a, \Gamma)\, \overline{AL}_i(a, \Gamma) \quad \text{for each } i \in N. \tag{10}$$

Making use of the notions of *brute good luck* and *action luck* we can now state that "*satisfaction* = *success* + *brute good luck*" and that "*success* = *power* + *action luck*", which yields that "*satisfaction* = *power* + *action luck* + *brute good luck*". The second and third equality only hold if we count weak swings and strong swings equally (i.e., if we take $\epsilon = 1$).[18] For the case $\epsilon = 1$ we shortly denote $\overline{POW}_i(a,\Gamma) = \overline{POW}_i^1(a,\Gamma)$ and $POW_i(\Gamma) = POW_i^1(\Gamma)$.

Proposition 4.12. *For a decision-making body applying an DMM* Γ *the following general equations hold:*

(i) $SAT(\Gamma) = SUC(\Gamma) + BGL(\Gamma)$
(ii) $SUC(\Gamma) = POW(\Gamma) + AL(\Gamma)$
(iii) $SAT(\Gamma) = POW(\Gamma) + AL(\Gamma) + BGL(\Gamma)$

Proof.

(i) For all $i \in N$, we have

$$\begin{aligned}
&SUC_i(\Gamma) + BGL_i(\Gamma)\\
&= \sum_{a\in\mathcal{A}_i^N} p(a,\Gamma)\, \overline{SUC}_i(a,\Gamma)\\
&\qquad + \sum_{a\in\mathcal{A}^N\setminus\mathcal{A}_i^N} p(a,\Gamma) \sum_{k\in K(a)} \tilde{p}(k,a,\Gamma)\overline{BGL}_i(k,a,\Gamma)\\
&= \sum_{\substack{a\in\mathcal{A}_i^N\\ a_i\sim o(a,\Gamma)}} p(a,\Gamma) + \sum_{a\in\mathcal{A}^N\setminus\mathcal{A}_i^N} p(a,\Gamma) \sum_{\substack{k\in K(a)\\ k_i\sim o(a,\Gamma)}} \tilde{p}(k,a,\Gamma)\\
&= \sum_{a\in\mathcal{A}^N} p(a,\Gamma) \sum_{\substack{k\in K(a)\\ k_i\sim o(a,\Gamma)}} \tilde{p}(k,a,\Gamma)\\
&= \sum_{a\in\mathcal{A}^N} p(a,\Gamma) \sum_{k\in K(a)} \tilde{p}(k,a,\Gamma)\, \overline{SAT}_i(k,\Gamma) = SAT_i(\Gamma),
\end{aligned}$$

where the first equality follows by definition of *SUC* and *BGL*, the second equality follows from the definitions of $\overline{SUC}$ and $\overline{BGL}$, the third equality follows since $k_i \sim a_i$ for all $a \in \mathcal{A}_i^N$ implies that

$$\sum_{\substack{a\in\mathcal{A}_i^N\\ a_i\sim o(a,\Gamma)}} p(a,\Gamma) = \sum_{a\in\mathcal{A}_i^N} \sum_{\substack{k\in K(a)\\ k_i\sim o(a,\Gamma)}} p(a,\Gamma)\, \tilde{p}(k,a,\Gamma),$$

[18] In fact, it is also possible to allow for any $\epsilon \in [0,1]$ but in that case we also need to redefine action luck taking account of weak and strong swings. Since this paper focusses on the distinction between success and satisfaction, we will not do that.

and the fourth and fifth equality follow from the definitions of $\overline{SAT}$ and SAT, respectively.

(ii) For all $i \in N$, we have

$$
\begin{aligned}
& POW_i(\Gamma) + AL_i(\Gamma) \\
&= \sum_{a \in \mathcal{A}_i^N} p(a,\Gamma)\, \overline{POW}_i(a,\Gamma) + \sum_{a \in \mathcal{A}_i^N} p(a,\Gamma)\, \overline{AL}_i(a,\Gamma) \\
&= \sum_{\substack{a \in \mathcal{A}_i^N \\ i \text{ has a swing in } a}} p(a,\Gamma) + \sum_{\substack{a \in \mathcal{A}_i^N,\ a_i \sim o(a,\Gamma) \\ i \text{ has no swing in } a}} p(a,\Gamma) \\
&= \sum_{\substack{a \in \mathcal{A}_i^N \\ i \text{ has a swing in } a \\ a_i \sim o(a,\Gamma)}} p(a,\Gamma) + \sum_{\substack{a \in \mathcal{A}_i^N \\ i \text{ has a swing in } a \\ a_i \not\sim o(a,\Gamma)}} p(a,\Gamma) \\
&\quad + \sum_{\substack{a \in \mathcal{A}_i^N,\ a_i \sim o(a,\Gamma) \\ i \text{ has no swing in } a}} p(a,\Gamma) \\
&= \sum_{\substack{a \in \mathcal{A}_i^N \\ i \text{ has a swing in } a \\ a_i \sim o(a,\Gamma)}} p(a,\Gamma) + \sum_{\substack{a \in \mathcal{A}_i^N,\ a_i \sim o(a,\Gamma) \\ i \text{ has no swing in } a}} p(a,\Gamma) \\
&= \sum_{\substack{a \in \mathcal{A}_i^N \\ a_i \sim o(a,\Gamma)}} p(a,\Gamma) = \sum_{a \in \mathcal{A}_i^N} p(a,\Gamma)\overline{SUC}_i(a,\Gamma) = SUC_i(\Gamma),
\end{aligned}
$$

where the first equality follows by definition of POW and AL, the second equality follows from the definitions of $\overline{POW}$ and $\overline{AL}$, the fourth equality follows since i has a swing in a implies that $a_i \sim o(a,\Gamma)$, and the fifth and sixth equality follow from the definitions of $\overline{SUC}$ and SUC, respectively.

(iii) obviously follows from (i) and (ii). □

If an DMM Γ applied by a decision-making body is a canonical one, we obtain the following general equation "*satisfaction* = *success* = *power* + *action luck*" which is quite close to Barry (1980), but contains the notion of *luck* as proposed by Laruelle and Valenciano (2005, 2008).

Corollary 4.13. *If the DMM Γ applied by a decision-making body is a canonical one, we obtain the following general equation:*

$$SAT(\Gamma) = SUC(\Gamma) = POW(\Gamma) + AL(\Gamma) \tag{11}$$

Proof. This follows since under the canonical set-up all actors act, and, thus, (1) *brute good luck* does not exist and (2) an actor has *success* if, and only if, it has *satisfaction*. □

Remark 4.14. Corollary 4.13 reflects the content of Remark 3.10, i.e., that *satisfaction* and *success* coincide under the canonical set-up.

What remains to be demonstrated is that there exist applications for which *brute good luck*, which creates the wedge between *satisfaction* and *success*, plays an essential role. This is done in Sect. 5 by two examples of a sequential DMM.

5 An Application: Sequential Decision-Making Mechanisms

Following van den Brink and Steffen (2008, 2012) let us assume a sequential "one desk" DMM. Regarding the nature of the actor we have to distinguish between three types of actors: (1) "bottom-", (2) "intermediate-", and (3) "top-actors". "*Bottom actors*" are those actors in the decision-making body who have a contact to the outside world and have the potential to receive new proposals, i.e., each decision-making body has at least one "bottom-actor". If there exists more than one bottom-actor we assume that a proposal enters the decision-making body via one of these actors with equal probability ("one desk" model). If a bottom-actor receives a proposal and does not support it by choosing the *no*-action, the proposal is regarded to be rejected by the decision-making body, i.e., the decision-making process is terminated. However, if the actor supports it by choosing the *yes*-action, it will be forwarded to the next actor in the decision-making process. This could be either an intermediate- or top-actor. It is an "*intermediate-actor*" if regarding the consequences of its choice of action the same applies as for a bottom-actor, i.e., the difference between both types of actors lies just in the fact whether the actor has a contact to the outside world and can receive new proposals or not. A "*top-actor*" is an actor who has no successors in the decision-making process.[19] Hence, its choice of action leads always to a final collective decision on the proposal, i.e., a rejection of the proposal if it chooses the "*no*-" and an approval if it chooses the *yes*-action. Hence, top-actors are the only actors in the decision-making body who can finally

[19]Note that the terminology successor-predecessor is opposite to the one as used in van den Brink and Steffen (2012). In van den Brink and Steffen (2012) both notions are used to refer to the positions of actors in a hierarchy, i.e., if actor i directly dominates an actor j, we say that i is a predecessor of j, and that j is a successor of i. In the present paper we make use of the same terminology to refer to actors in a sequential DMM, i.e., if an actor i chooses its action after actor j has made its choice of action, we say that actor i is a successor of j, and that j is a predecessor of i.

enforce an approval of a proposal, while all actors have the ability to reject it (when they are allowed to choose an action). This implies that any sequential DMM in addition to at least one bottom-actor must also contain at least one top-actor, while the existence of intermediate-actors is not a necessary requirement.

Assumption 5.1. *We add the following assumptions to those made under Assumption 2.1, where assumption 5 is now replaced by assumption 7:*

7. *All actors choose their action sequentially, where each actor $i \in N$ has not more than one chance to be involved in the decision-making, i.e., to choose its action.*
8. *New proposals entering the decision-making body can only be received by the bottom-actor(s) being actors in positions with no predecessors in the decision-making process. This process continues then via the intermediate-actor(s)—if they exist—to the top-actor(s).*
9. *A new proposal can only be received by one bottom-actor at the same time (One Desk Model).*
10. *The choice of the* yes-*action results (1) in a final approval if actor i is the top-actor, i.e. if i has no successor in the decision-making process, or (2) in forwarding the proposal to one or more successors if i is not the top-actor.*
11. *The choice of the* no-*action results in a final rejection of the proposal, if (1) actor i is a bottom actor, or (2) if for the actor who has forwarded the proposal to i there is no other successor in the decision-making process left to ask for supporting the proposal whose individual support contains the potential of a final approval. If such other successor as in case (2) exists it results in forwarding the proposal to this actor.*

Remark 5.2. Note that under DMMs applying in addition to Assumption 2.1 also Assumption 5.1—the latter contains a set of assumptions which are characteristic for hierarchical organizations (see van den Brink and Steffen 2008, 2012)—an actor i can be excluded from the decision-making for two reasons: (1) if i is an intermediate- or top-actor the decision-making process could have already been terminated by another actor (which could be a bottom- or intermediate-actor) by choosing the *no*-action, or (2) if the decision-making body contains more than one bottom-actor and the proposal has entered the decision-making body via a bottom-actor such that subsequent decision-making process does not include i as an actor.

Clearly, in such sequential DMM's actors who are non-top-actors have *strong* and *weak swings*. A non-top-actor has a strong swing in action profiles where it is allowed to choose an action and where it chooses the *yes*-action and all its successors in the decision-making process choose the *yes*-action as well (leading to an approval of the proposal), while the proposal would be rejected if the actor chooses the *no*-action. It has a *weak swing* in every action profile where a non-top-actor is allowed to choose an action and where it chooses the *no*-action (leading to rejection of the

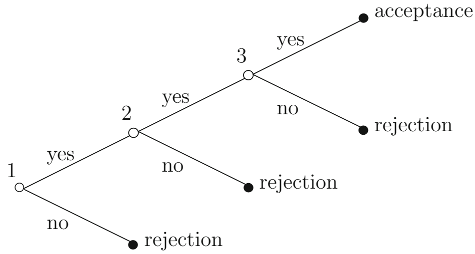


Fig. 1 Extensive game form Γ_1

proposal), while the proposal could be approved (if all its successors in the decision-making process who are allowed to choose an action choose the *yes*-action), but also could still be rejected (if there is at least one successor who is allowed to choose an action and chooses the *no*-action). A top-actor has two swings which are both *strong*.

Making use of Assumptions 2.1 and 5.1 we will now investigate two simple sequential DMMs in order to illustrate the relationship between *satisfaction*, *success*, and *power* and the usage of the related measures.[20]

Example 5.3. Let Γ_1 be a sequential DMM with $N = \{1, 2, 3\}$ where all actors choose their action sequentially in numerical ascending order, i.e., actor 1 is the unique bottom-, actor 2 is an intermediate-, and actor 3 is the unique top-actor. Γ_1 can be represented by the extensive game form as given by Fig. 1. Our *satisfaction*, *success*, and *power* analysis is displayed by Table 2 assuming uniform probability distributions.

As it can be seen from Fig. 1 and Table 2 under Γ_1 we have four action profiles a. In one of them actor 3 and in another one actors 2 and 3 are excluded from the decision-making due to the fact that another actor, i.e., actor 2 in the former and actor 1 in the latter case, has chosen the *no*-action. *Satisfaction* of the excluded actors under these profiles is then a result of their *brute good luck*, only.

The next example includes also the second type of exclusion as explained in Remark 5.2, i.e., exclusion due to the existence of more than one bottom-actor.

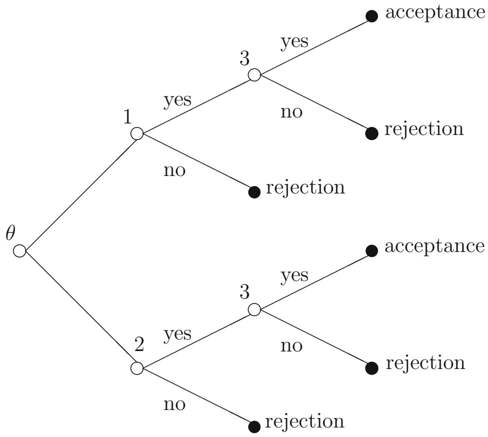
Example 5.4. Let Γ_2 be a sequential DMM with $N = \{1, 2, 3\}$. Let us assume that nature chooses with probability 0.5 whether bottom-actor 1 or 2 will receive a new proposal. After having received a proposal the bottom-actors can reject a proposal on their own, while for an approval both require the consent of top-actor 3. Γ_2 can

[20]Note that Γ_1 and Γ_2 in Examples 5.3 and 5.4, respectively, could be regarded as examples for DMMs in hierarchical organizations, where the structure of the hierarchy is a "line" in case of Γ_1 and a "star" in case of Γ_2 (see van den Brink and Steffen 2008, 2012).

Table 2 Satisfaction, success, and power in Γ_1

	a					a-compatible k			$\overline{SAT}_i$			$\overline{SUC}_i$			$\overline{POW}_i^{\epsilon}$		
$p(a, \Gamma_1)$	1	2	3	o	$p(\mathrm{k}, a, \Gamma_1)$	1	2	3	1	2	3	1	2	3	1	2	3
0.25	*Yes*	*Yes*	*Yes*	*Approval*	1.00	*Approve*	*Approve*	*Approve*	1	1	1	1	1	1	1	1	1
0.25	*Yes*	*Yes*	*No*	*Rejection*	1.00	*Approve*	*Approve*	*Reject*			1			1			1
0.25	*Yes*	*No*		*Rejection*	0.50	*Approve*	*Reject*	*Approve*		1			1			ε	
				Rejection	0.50	*Approve*	*Reject*	*Reject*		1	1						
0.25	*No*			*Rejection*	0.25	*Reject*	*Approve*	*Approve*	1			1			ε		
				Rejection	0.25	*Reject*	*Reject*	*Approve*	1	1							
				Rejection	0.25	*Reject*	*Approve*	*Reject*	1		1						
				Rejection	0.25	*Reject*	*Reject*	*Reject*	1	1	1						
						Measure			**0.50**	**0.63**	**0.75**	**0.50**	**0.50**	**0.50**	**0.25+ 0.25ε**	**0.25+ 0.25ε**	**0.50**

Fig. 2 Extensive game form Γ_2



be represented by the extensive game form as given by Fig. 2. Our *satisfaction*, *success*, and *power* analysis is displayed by Table 3 assuming uniform probability distributions.

6 Concluding Remarks

The major purpose of this paper is to disentangle the relationship between *satisfaction* and *success*. Despite the fact that both notions are conceptually distinct they are mixed-up in the voting power literature. We have pointed out that a potential explanation for this phenomenon is the fact, that both notions coincide under the canonical set-up which assumes a simultaneous DMM. Clarifying the relationship between both notions we found that the notion of *luck* requires an disentanglement as well. We illustrated the requirement for the disentanglements of all three notions by replacing the canonical simultaneous DMM by a sequential DMM, which allows that, under specific circumstances, some actors are excluded from the decision-making.

Now we would like to wind-up this paper with some remarks on the relationship between *satisfaction* and *success* if we allow for "abstention" as a *tertium quid*. That is what Barry (1980) assumes when he illustrates his notion of *luck* under a simultaneous set-up. Hence, one might be tempted to argue that *brute good luck* can also exist under a simultaneous DMM if one allows for "abstention" as a *tertium quid*. However, this would be mistaken as allowing for "abstention" is nothing else than an extension of the action set: "abstention" is an action an actor can choose in addition to the "*yes*-" and "*no*-action". In contrast to the "*non*-action" of an excluded actor under a sequential DMM, an actor choosing the "*abstention*-action" is not excluded from the decision-making by the definition of the decision-making

Table 3 Satisfaction, success, and power in Γ_2

	a					*a-compatible k*			$\overline{SAT}_i$			$\overline{SUC}_i$			$\overline{POW}_i^{\epsilon}$		
$p(a, \Gamma_2)$	1	2	3	*o*	$p(\text{k},a, \Gamma_2)$	1	2	3	1	2	3	1	2	3	1	2	3
0.50×0.33	*Yes*		*Yes*	*Approval*	0.50	*Approve*	*Approve*	*Approve*	1	1	1	1		1	1		1
					0.50	*Approve*	*Reject*	*Approve*	1		1						
0.50×0.33	*Yes*		*No*	*Rejection*	0.50	*Approve*	*Approve*	*Reject*			1			1			1
					0.50	*Approve*	*Reject*	*Reject*		1	1						
0.50×0.33	*No*			*Rejection*	0.25	*Reject*	*Approve*	*Approve*	1			1			ε		
					0.25	*Reject*	*Approve*	*Reject*	1		1						
					0.25	*Reject*	*Reject*	*Approve*	1	1							
					0.25	*Reject*	*Reject*	*Reject*	1	1	1						
0.50×0.33		*Yes*	*Yes*	*Approval*	0.50	*Approve*	*Approve*	*Approve*	1	1	1		1	1		1	1
					0.50	*Reject*	*Approve*	*Approve*		1	1						
0.50×0.33		*Yes*	*No*	*Rejection*	0.50	*Approve*	*Approve*	*Reject*			1			1			1
					0.50	*Reject*	*Approve*	*Reject*	1		1						
0.50×0.33		*No*		*Rejection*	0.25	*Approve*	*Reject*	*Approve*		1			1			ε	
					0.25	*Approve*	*Reject*	*Reject*		1	1						
					0.25	*Reject*	*Reject*	*Approve*	1	1							
					0.25	*Reject*	*Reject*	*Reject*	1	1	1						
						Measure			**0.58**	**0.58**	**0.83**	**0.33**	**0.33**	**0.67**	**0.17+ 0.17ε**	**0.17+ 0.17ε**	**0.67**

procedure, but only due to its own choice of action not to opt in favor or against a proposal. Hence, we cannot ascribe *brute good luck* to such an actor. However, we also cannot ascribe *action luck* to such an actor as its choice of action does not correspond to the elements of the binary outcome set. What we require is a third notion of *luck* which we may call *abstention luck* being an additional component of *satisfaction*: "*satisfaction* = *success* + *brute good luck* + *abstention luck*".

Regarding the ascription of *abstention luck* we have to distinguish between two principal cases if an actor chooses the "*abstention*-action": (1) the actor has an inclination for one of the two elements of the outcome set, i.e., either to approve or to reject the proposal, but for certain reasons decides to choose the "*abstention*-action", for instance, because the costs for performing the other actions are too high, or (2) the actor is indifferent between the two elements of the outcome set. In the latter case, given Assumption 2.1, we will always ascribe *satisfaction* to the actor with respect to the collective outcome, while in the former case *satisfaction* will only be ascribed if the actor's specific inclination for a unique element of the outcome set corresponds to the collective outcome.

Finally, by making use of Barry's (1980) terminology we would like to point out that the difference between *success* and *abstention luck* can be characterized by the fact whether an actor "has tried" or "has not tried". Remember that in Sect. 1 we noted that for his definition of *success* Barry (1980) adds the requirement that "the actor has tried" which we used as a synonym for "having chosen an action". While both coincide under the canonical and our sequential set-up (see Footnote 3), this no longer holds if we allow for "abstention" as a *tertium quid*. In this case we have to distinguish between two types of actions: (1) actions which imply that by their choice an actor "tries" to obtain a specific collective outcome, e.g., the approval of a proposal by choosing the "*yes*-action", and (2) actions which imply that by their choice an actor "does not try" to obtain a specific outcome. Actions of type (1) can lead to *success*, while actions of type (2) can lead to *abstention luck*.

References

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

Barry, B. (1980). Is it better to be powerful or lucky? *Political Studies*, *28*, 183–194, 338–352.

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

Braham, M., & Steffen, F. (2002). Voting power in games with abstentions. In M. J. Holler, H. Kliemt, D. Schmidtchen, & M. E. Streit (Eds.) *Power and fairness* (pp. 333–348). Tübingen: Mohr-Siebeck.

Brams, S. J., & Lake, M. (1978). Power and satisfaction in a representative democracy. In P. C. Ordeshook (Ed.), *Game theory and political science* (pp. 529–562). New York: New York University Press.

van den Brink, R., Rusinowska, A., & Steffen, F. (2011). Measuring power and satisfaction in societies with opinion leaders: dictator and new opinion leader properties. *Homo Oeconomicus*, *28*, 161–185.

van den Brink, R., Rusinowska, A., & Steffen, F. (2013). Measuring power and satisfaction in societies with opinion leaders: an axiomatization. *Social Choice and Welfare, 41*, 671–683.
van den Brink, R., & Steffen, F. (2008). Positional power in hierarchies. In M. Braham, & F. Steffen (Eds.), *Power, freedom, and voting* (pp. 57–81). Berlin, Heidelberg: Springer.
van den Brink, R., & Steffen, F. (2012). Axiomatizations of a positional power score and measure for hierarchies. *Public Choice*, *151*, 757–787.
Cohen, G. A. (1989). On the currency of egalitarian justice. *Ethics*, *99*, 906–944.
CollinsDictionary (2009) *Collins English Dictionary* , 10th edn. Glasgow: HarperCollins.
Dowding, K. (1991). *Rational choice and political power*. Aldershot: Edward Elgar.
Dowding, K. (1996). *Power*. Buckingham: Open University Press.
Dubey, P., & Shapley, L. S. (1979). Mathematical properties of the Banzhaf power index. *Mathematics of Operations Research*, *4*, 99–131.
Dworkin, R. (1981). What is equality? Part 2: equality of resources. *Philosophy of Public Affairs*, *10*, 283–345.
Felsenthal, D., & Machover, M. (1997). Ternary voting games. *International Journal of Game Theory*, *26*, 335–351.
Felsenthal, D., & Machover, M. (1998). *The Measurement of voting power: theory and practice, problems and paradoxes*. London: Edward Elgar.
Felsenthal, D., & Machover, M. (2001). Models and reality: the curious case of the absent abstention. In M. J. Holler, & G. Owen (Eds.), *Power indices and coalition formation* (pp. 87–103). Dordrecht: Kluwer.
Grabisch, M., & Rusinowska, A. (2010). A model of influence in a social network. *Theory and Decision*, *69*, 69–96.
Holler, M. J. (1982). Forming coalitions and measuring voting power. *Political Studies*, *30*, 262–271.
Hosli, M. O., & Machover, M. (2004). The nice treaty and voting rules in the council: a reply to Moberg (2002). *Journal of Common Market Studies*, *42*, 497–521.
Laruelle, A., Martinez, R., & Valenciano, F. (2006). Success versus decisiveness: conceptual discussion and case study. *Journal of Theoretical Politics*, *18*, 185–205.
Laruelle, A., & Valenciano, F. (2005). Assessing success and decisiveness in voting situations. *Social Choice and Welfare*, *24*, 171–197.
Laruelle, A., & Valenciano, F. (2008). *Voting and collective decision-making*. Cambridge: Cambridge University Press.
Morriss, P. (1987/2002). *Power: a philosophical analysis*. Manchester: Manchester University Press.
Nevison, C. H. (1979). Structural power and satisfaction in simple games. In S. J. Brams, et al. (Eds.), *Applied game theory* (pp. 39–57). Würzburg: Physica.
Nevison, C. H. (1982). Structural satisfaction in simple games. *Mathematical Social Sciences*, *3*, 397–401.
Penrose, L. S. (1946). The elementary statistics of majority voting. *Journal of the Royal Statistical Society*, *109*, 53–57.
Pinker, S., et al. (Eds.). (2011) *The American heritage dictionary of the English language*, 5th edn. Boston: Houghton Mifflin Harcourt.
Rae, D. (1969). Decision rules and individual values in constitutional choice. *American Political Science Review*, *63*, 40–56.
Roemer, J. (1986). Equality of resources implies equality of welfare. *Quarterly Journal of Economics*, *101*, 751–784.
Rusinowska, A., & de Swart, H. (2006). Generalizing and modifying the Hoede-Bakker index. In H. de Swart et al. (Eds.), *Theory and applications of relational structures as knowledge instruments*. Springer's Lecture Notes in Artificial Intelligence, LNAI 4342 (pp. 60–88). Heidelberg: Springer.
Straffin, P. D. (1977). Majority rule and general decision rules. *Theory and Decision*, *8*, 351–360.

Straffin, P. D. (1978). Probability models for power indices. In P. C. Ordeshook (Ed.), *Game theory and political science* (pp. 477–510). New York: New York University Press.

Straffin, P. D., Davis, M. D., & Brams, S.J. (1982). Power and satisfaction in an ideologically divided voting body. In M. J. Holler (Ed.) *Power, voting and voting power* (pp. 239–255). Würzburg: Physica.

Voting Power Techniques: What Do They Measure?

Sreejith Das

1 Introduction

Voting power is a field of co-operative game theory that has seen a recent resurgence, due, in no small part, to the work of Felsenthal and Machover (1998) and their seminal book. Despite the importance of the field, it is a subject that is not studied widely enough, and is poorly understood outside of the voting power community.

The concept behind voting power is simple enough. The idea is to measure the ability of an individual voter to affect the outcome of a voting game. This kind of analysis is invaluable when it comes to designing fair, and democratic, institutions. For instance, most people would agree that it is desirable to design voting within the European Union such that a country with twice the population should have twice the influence, compared with a country half the size. But the question remains, how do you go about measuring voting power?

In the literature, there have been a number of techniques proposed to measure voting power, such as Shapley and Shubik (1954), Banzhaf (1965), Coleman (1971), Deegan and Packel (1978), Johnston (1978), Straffin (1977). The two most widely used techniques are by Banzhaf, and Shapley and Shubik. It has previously

The author would like to thank Maurice Salles and Sean Breslin for their help in the preparation of this manuscript.

S. Das (✉)
Department of Economics, Mathematics and Statistics, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK
e-mail: sreejith.das@gmail.com

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare, DOI 10.1007/978-3-319-05158-1_5,

been proposed, by Straffin (1977, 1978), that the differences between these two techniques rest solely in the underlying probability models. However, this paper will show that the differences are much more fundamental.

At a recent conference in London, Dan Felsenthal explained that while many of us were proficient in using voting power techniques, he argued that no one really understood what these techniques were actually calculating. But he hoped that one day, we would. At the same conference, Moshé Machover suggested that adopting a more probabilistic approach to voting power would be beneficial. It is hoped that this work will go some way towards satisfying both of their aspirations.

The paper is organised as follows: Section 2 introduces a new methodology that will be employed to analyse the standard voting power techniques. Section 3 examines the standard techniques, and re-interprets them with the new methodology. Section 4 analyses the standard techniques when applied to simple voting games, and derives what each technique is actually calculating. Section 5 summarises the main contributions of this paper.

The results presented in the main body of the paper are generalised in Appendix 1 to encompass games in which voters can do more than simply vote "yes" or "no". Appendix 2 extends the analysis to encompass games with multiple possible outcomes, and arbitrarily complex decision rules. Finally, Appendix 3 gives the rigorous mathematical definitions used in the preceding analyses.

This paper attempts to present measure theoretic ideas to as wide an audience as possible. On occasion, this will result in some mathematical notation being over-simplified for ease of comprehension. Any mathematically inclined readers are asked to forgive these compromises, and make use of Appendix 3 instead.

2 Counting Blocks

Now for a short digression from voting power theory. Imagine you work for a Danish toy manufacturer of children's interlocking building blocks, and it has just started a recycling scheme. The amount of money it is willing to pay for a batch of blocks is dependent upon the percentage of blue blocks in the shipment (for some reason, the blue blocks are more valuable).

Now imagine you've just been handed a large pile of blocks, which we will call Ω. It's your job to calculate the percentage of blue blocks in the batch. Being an industrious type, you decide to build a machine to do this for you.

The first stage in your plan is to count how many blocks there are in total. Let's call this block counting machine $\mathbb{P}$. After one run through, we'll know how many blocks we have in Ω, we'll call this number $\mathbb{P}(\Omega)$. Your new machine will output something like $\mathbb{P}(\Omega) = 1{,}034$, or $\mathbb{P}(\Omega) = 32$, depending on how many blocks there are in the batch (Fig. 1).

With amazing forethought you realise that, as you need to calculate a percentage, it will be more useful to have $\mathbb{P}(\Omega) = 100\,\%$ after all the blocks have been counted.

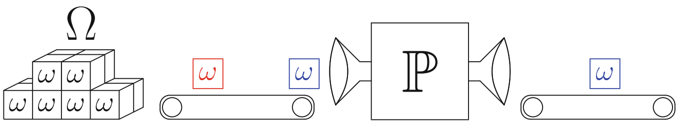


Fig. 1 A block counting machine

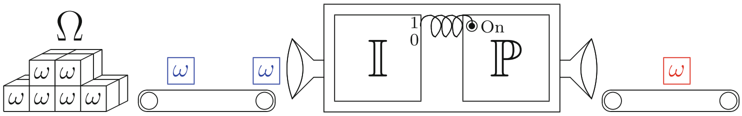


Fig. 2 A super-counter machine

You adjust the machine so that instead of adding 1 every time a little ω goes past it will add $\frac{1}{|\Omega|}$.[1] Now, after all the blocks have passed through, the machine will read $\mathbb{P}(\Omega) = 1$ (which is, of course, equivalent to $\mathbb{P}(\Omega) = 100\,\%$).

The second stage in your plan is to add a "magic eye" machine that can "see" if a blue block has gone past, we'll call this the $\mathbb{I}$ machine. The $\mathbb{I}$ machine is very basic, it simply outputs $\mathbb{I}(\omega) = 1$ if it sees a blue block, and $\mathbb{I}(\omega) = 0$ otherwise.

The final stage in your plan is to link the "magic eye" machine with the block counting machine, to create a super-counter machine. You connect the output of the $\mathbb{I}$ machine, with the "on/off" switch of the $\mathbb{P}$ machine. Now, whenever a blue block goes by, the $\mathbb{I}$ machine will turn on the $\mathbb{P}$ machine, allowing it to count. But if a non-blue block should pass, the $\mathbb{I}$ machine will turn off the $\mathbb{P}$ machine, preventing it from counting (Fig. 2).

And that's it! The combined $\mathbb{I}$ and $\mathbb{P}$ machines work together to calculate the percentage of blue blocks. After all the blocks have gone through the super-counter machine, the output of the $\mathbb{P}$ machine will be the percentage of blocks that are blue.

The operation of the super-counter can be described as follows:

(1) Start with a pile of blocks called Ω.
(2) Take each little block ω in turn, and send it through the super-counter.
(3) If ω is blue, the $\mathbb{I}$ machine will turn on the $\mathbb{P}$ machine.
(4) If ω is not blue, the $\mathbb{I}$ machine will turn off the $\mathbb{P}$ machine.
(5) The result is given by reading the output of the $\mathbb{P}$ machine after all the blocks have passed through the super-counter.

[1]For the purposes of this example, we shall ignore how you can come to know $\frac{1}{|\Omega|}$ before all the blocks have been counted!

2.1 The Maths

As a bit of a mathematician, you want to write down the operation of the super-counter using mathematical notation. Let's start by writing down what happens when a single block passes through the machine. We can mimic the action of the $\mathbb{I}$ machine turning the $\mathbb{P}$ machine on and off by multiplying $\mathbb{I}$ and $\mathbb{P}$ together (remember that the $\mathbb{I}$ machine outputs 1 if it is blue, and 0 otherwise).

$$\mathbb{I}(\omega) \times \mathbb{P}(\omega).$$

Next we have to represent every little block ω moving through the machine, with the result added to a running count. We could use the $\sum$ notation for this, but, for our purposes, the integral notation would be better.

$$\int_{\omega \in \Omega} \mathbb{I}(\omega) \times \mathbb{P}(\omega).$$

This integral notation simply says, take every single little block ω from the big pile of blocks Ω, and send it through the $\mathbb{I}$ and $\mathbb{P}$ machines.

We're almost done, just a few more tweaks. First, let's rename the $\mathbb{I}$ function to $\mathbb{I}^{Blue}$, because the $\mathbb{I}$ machine is looking for blue blocks. Second, we get rid of the redundant $\times$ sign between $\mathbb{I}$ and $\mathbb{P}$. And third, in keeping with standard notation, we change the final ω to $d\omega$.[2]

$$\int_{\omega \in \Omega} \mathbb{I}^{Blue}(\omega)\ \mathbb{P}(d\omega).$$

We finish off our mathematical expression of the super-counter by writing down what this super-counter was designed to do. Which, in this case, is to calculate the percentage of blocks that are blue. Mathematically speaking, we can call this the probability of a block being blue.

$$\Pr(Blue) = \int_{\omega \in \Omega} \mathbb{I}^{Blue}(\omega)\ \mathbb{P}(d\omega).$$

2.2 Non-uniform Blocks

Satisfied in your new super-counter machine, you patiently wait for your first batch of blocks to arrive. When they finally do, you receive an unwelcome surprise. Instead of a nice neat pile of individual blocks, you are given a huge mess of blocks

---

[2]It is customary to write $\int_x y(x)\ dx$ instead of $\int_x y(x)\ x$.

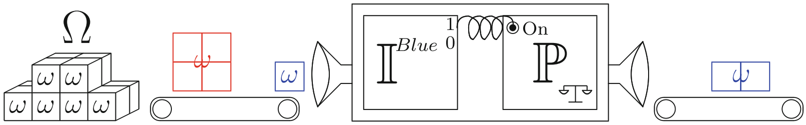


Fig. 3 A super-measurer

stuck together in clumps of different sizes. The blocks have come from a school maths department where they were using them to illustrate factorials. The blocks have arrived in clumps of size 1!, 2!, 3!, 4!, and so on. As luck would have it, each clump is made up of one colour only. Despite this, before you can use your machine, you'll have to break up these clumps into their individual little blocks. If only there was some way to modify the super-counter to cope with these clumps automatically? Fortunately, there is. And it's all to do with the $\mathbb{P}$ machine.

Instead of using the $\mathbb{P}$ machine to count blocks as they go past, the $\mathbb{P}$ machine can weigh them instead. This simple change means that even if a clump of x blocks were to go through the machine it would still know how many went past, because they would weigh x times as much as an individual block (Fig. 3).

As we're not really counting anymore, the machine should be renamed. It could be called a super-weigher, but calling it a super-measurer would be even better. This new super-measurer works as follows. (In our previous example, we used ω to represent an individual block, this time we can use it to represent a clump.)

(1) Start with a pile of blocks called Ω.
(2) Take each clump ω in turn, and send it through the super-measurer.
(3) If the clump is blue, use the $\mathbb{I}$ machine to turn on the weighing machine $\mathbb{P}$.
(4) If the clump is not blue, use the $\mathbb{I}$ machine to turn off the weighing machine $\mathbb{P}$.
(5) The result is given by reading the total weight measured by $\mathbb{P}$ after all clumps have passed through the super-measurer.

Expressing the operation of the super-measurer mathematically gives,

$$\Pr(Blue) = \int_{\omega \in \Omega} \mathbb{I}^{Blue}(\omega)\ \mathbb{P}(d\omega).$$

This is the same mathematical representation as the super-counter machine! How can this even be possible? It's because the weighing of the clumps has been incorporated into the $\mathbb{P}$ function. Mathematically, we say that $\mathbb{P}$ is a measure on the subsets of Ω, and $\mathbb{I}$ is an indicator function.

2.3 Discussion

Let's examine the super-measurer in greater detail. It calculates the probability of a block being blue, no matter how weird and clumpy the set Ω might be. It could be made up of uniform 1 block clumps, or they could be some weird number like $(|C|-1)!(|N|-|C|)!$ in size. The super-measurer doesn't even care in which order the clumps pass through, it will still calculate Pr(*Blue*) in the end.

In other words, *changing the distribution of the blocks doesn't change the statistic being calculated.*

3 The Established Voting Power Calculation Methods and Their Indicator Functions

Let's remind ourselves how we go about measuring something.

$$\Pr(Something) = \int_{\omega \in \Omega} \mathbb{I}^{Something}(\omega)\ \mathbb{P}(d\omega).$$

We need an indicator function that identifies the property we want to measure, and we need a special $\mathbb{P}$ "weighing" function, defined on the subsets of Ω, to add things up. Of these two, only the indicator defines the statistic being calculated, or measured. So, if we want to understand what the standard voting power calculation techniques are actually calculating, we need only understand their indicator functions.

3.1 Shapley-Shubik Technique

Shapley and Shubik (1954) state that the power of an individual member of a legislative body depends on the chance they have of being critical to the success of a winning coalition. They explain that a voter can be "pivotal" when they can turn a possible defeat into a success. And they construct their index as follows.

(1) There is a group of individuals willing to vote for a bill.
(2) They vote in order.
(3) As soon as a majority has voted for it, it is declared passed.
(4) The (pivotal) member who voted last is given credit for passing the bill.

The voting orders are chosen randomly, and they calculate the number of times that a voter is considered pivotal. The final Shapley-Shubik index is produced by dividing the pivotal count by the total number of voting orders (i.e. $n!$, where n is the number of voters). They describe this as the frequency with which a particular voter is considered "pivotal".

For a moment, let's examine their term pivotal. It requires a losing voting scenario in which the voter initially votes "no" to become winning when they vote "yes" instead. Rather than call the voter pivotal, let's call it critical instead. Furthermore, as the voter becomes critical by increasing its support, let's call it increasingly critical. Finally, as the pivotal voter always starts off by expressing zero support (voting "no") for the bill, we will call it Increasing Criticality 0. Hence, the Shapley-Shubik index is given by the following algorithm.

(1) Examine every possible voting order.
(2) For each voting order identify if it is Increasing Criticality 0 for the given voter.
(3) If so, add 1 to a running count for the given voter.
(4) Repeat until all voting orders have been examined, then divide by $n!$.

It is explicit within the construction of the Shapley-Shubik index that all voting orders are equiprobable, the term $\frac{1}{n!}$ is the probability of a given voting order arising in a game with n voters. With this is mind, it is easy to see that the Shapley-Shubik index is nothing more than the probability of a voter being Increasing Criticality 0. If we let ω represent a voting order, then we could create a super-measurer to calculate the Shapley-Shubik index as follows,

$$\text{ShapleyShubik}_i = \int_\omega \mathbb{I}^{IC_i^0}(\omega)\ \mathbb{P}(d\omega) = \Pr(IC_i^0),$$

where $\mathbb{P}(d\omega) = \frac{1}{n!}$, and:

$$\mathbb{I}^{IC_i^0}(\omega) = \begin{cases} 1 \text{ if } \omega \text{ is Increasing Criticality 0 for the given voter } i; \\ 0 \text{ otherwise.} \end{cases}$$

3.2 *Banzhaf Technique*

Banzhaf (1965) states that power in a legislative sense is the ability to affect outcomes. He says the power of a legislator is given by the number of possible voting combinations of the entire legislature in which the legislator can alter the outcome by changing their vote. The key point to understand with the Banzhaf method is that it doesn't restrict itself in anyway when it comes to identifying critical voting combinations. It doesn't care if the voting combination is initially losing, or initially winning. It doesn't care if the voter was initially voting "yes" or initially voting "no". It just counts up the maximum number of possible situations in which the voter can change the outcome by changing its vote.

We can interpret the ability to alter the outcome through a change of vote as follows: a voter is able to make a losing outcome winning by increasing their support (which we've previously termed Increasing Criticality), or a voter is able to make a winning outcome losing by decreasing their support (which we will call Decreasing Criticality). We can name the combination of both Increasing and Decreasing Criticality as Total Criticality.

Given that Banzhaf measures every situation in which a voter can be critical, it must be a Total Criticality measure. Furthermore, as Banzhaf makes no specific requirement for the voter to be initially voting one way or the other, we will call this a Total Criticality δ measure, where the symbol δ shows that we don't care how the voter was initially voting. Hence, the Banzhaf voting power method is given by the following algorithm.

(1) Examine every possible voting combination.
(2) For each voting combination identify if it is Total Criticality δ for the given voter.
(3) If so, add 1 to a running count for the given voter.
(4) Repeat until all voting combinations have been examined, then divide by 2^n.

Banzhaf assumes that every voting combination is equiprobable, the term $\frac{1}{2^n}$ is the probability of a given voting combination arising (where n is the number of voters). With this is mind, it is easy to see that the Banzhaf measure is nothing more than the probability of a voter being Total Criticality δ. If we use the symbol ω to represent a voting combination, then we can create a super-measurer to calculate the Banzhaf measure as follows,

$$\text{Banzhaf}_i = \int_{\omega} \mathbb{I}^{TC_i^{\delta}}(\omega)\ \mathbb{P}(d\omega) = \Pr(TC_i^{\delta}),$$

where $\mathbb{P}(d\omega) = \frac{1}{2^n}$ and,

$$\mathbb{I}^{TC_i^{\delta}}(\omega) = \begin{cases} 1 \text{ if } \omega \text{ is Total Criticality } \delta \text{ for the given voter } i\,; \\ 0 \text{ otherwise.} \end{cases}$$

3.3 Straffin

Straffin (1977) proposed two different techniques differentiated by the probability model assumed. The Independence Assumption technique assumes all voters vote in favour with a common probability $p = 0.5$. The Homogeneity Assumption also makes use of a common probability p, but p is allowed to vary uniformly between [0, 1] by integrating it over the aforementioned range.

Straffin defines his measure as the probability that voter i's vote will make a difference in the outcome. Making it, like Banzhaf, a measure of Total Criticality. And, as there is no requirement for voter i to be initially voting one way or another, it is a measure of Total Criticality δ.

$$\text{Straffin}_i = \int_{\omega} \mathbb{I}^{TC_i^{\delta}}(\omega)\ \mathbb{P}(d\omega) = \Pr(TC_i^{\delta}).$$

Both the Independence Assumption technique, and the Homogeneity Assumption technique are given by $\Pr(TC_i^\delta)$. The different probability models are absorbed by the $\mathbb{P}(d\omega)$ term.

3.4 Coleman

Of all the researchers working in the field of voting power theory, Coleman (1971) was perhaps the first to appreciate the subtle differences that exist between Increasing and Decreasing criticality.[3] He defined two measures of power, the power to initiate action, and the power to prevent action.

The initiate action measure is a count of the number of times a voter can be critical to a losing coalition, divided by the number of losing coalitions. While the prevent action measure is a count of the number of times a voter can be critical to a winning coalition, divided by the number of winning coalitions.

$$\text{Coleman Initiate Action}_i = \frac{\int_\omega \ \mathbb{I}^{IC_i^\delta}(\omega) \ \ \mathbb{P}(d\omega)}{\int_\omega \ \mathbb{I}^{\text{Losing}}(\omega) \ \ \mathbb{P}(d\omega)} = \frac{\Pr(IC_i^\delta)}{\Pr(\text{Losing})}.$$

$$\text{Coleman Prevent Action}_i = \frac{\int_\omega \ \mathbb{I}^{DC_i^\delta}(\omega) \ \ \mathbb{P}(d\omega)}{\int_\omega \ \mathbb{I}^{\text{Winning}}(\omega) \ \ \mathbb{P}(d\omega)} = \frac{\Pr(DC_i^\delta)}{\Pr(\text{Winning})}.$$

3.5 Johnston

The Johnston (1978) index can be described as follows. Examine every winning coalition, identify those members which can destroy it, and then allocate a point, or fraction of a point, to them. In other words, this is a measure of Decreasing Criticality. From his paper, it seems reasonable to assume that his index requires the voter to express zero approval in order to destroy the coalition, so we will call it a Decreasing Criticality 0 measure.

$$\text{Johnston}_i = \int_\omega \ \mathbb{I}^{DC_i^0}(\omega) \ \ \mathbb{P}(d\omega) \ = \ \Pr(DC_i^0).$$

Both the original version of the Johnston index (where one point is added for every destroyable coalition), and the modified version (where a fraction of a point is added) are given by $\Pr(DC^0)$. In the modified version, the fraction that is added is

[3] Shapley and Shubik (1954) understood it was possible to be Decreasingly Critical, but they did not appreciate that this was materially different to being Increasingly Critical (examine their comments regarding their proposed "blocking index").

a function of ω only, hence it can be absorbed within the $\mathbb{P}(d\omega)$ function. Ergo, the modified version is the same as the original version, albeit with a different probability model.[4]

3.6 Deegan-Packel and Holler Public Good Indices

The Deegan and Packel (1978), and Holler (1982) indices are incredibly similar. Both indices can be described by the following.

(1) Examine every minimum winning coalition.
(2) Identify if it is Decreasingly Critical 0 for the given voter.
(3) If so, add 1 (for Holler), or a fraction of 1 (for Deegan-Packel) to a running count for the given voter.
(4) Repeat until all minimum winning coalitions have been examined.

In the Deegan-Packel index, the fraction that is added is a function of ω only, hence it can be absorbed within the $\mathbb{P}(d\omega)$ function of the game. In other words, both the Deegan-Packel and Holler Public Good indices are the same, albeit with slightly different probability models (in the same way that the Johnston and modified Johnston indices are the same).[5] They are both some kind of Decreasing Criticality 0 measure.

Let's focus upon the Deegan-Packel index, we note, from their paper, that their probability model assumes that only minimum winning coalitions (MWC) will form.[6] Therefore, the probability model implicit within these indices ensure that $\mathbb{P}(d\omega) = 0$ unless $\omega \in MWC$. In other words, the $\mathbb{P}$ functions ensure that an integration over the set of minimum winning coalitions is equivalent to an integration over the entire set Ω. Hence,

$$\text{DeeganPackel}_i = \int_{\omega\in\Omega} \mathbb{I}^{DC_i^\delta}(\omega)\ \mathbb{P}(d\omega) = \Pr(DC_i^0).$$

$$\text{HollerPGI}_i = \int_{\omega\in\Omega} \mathbb{I}^{DC_i^\delta}(\omega)\ \mathbb{P}(d\omega) = \Pr(DC_i^0).$$

[4] The actual fraction that is added is inversely proportional to the number of voters that express full support in ω. Hence, the probability model of the modified index implies that coalitions with more voters expressing full support are less likely to occur.

[5] The actual fraction that is added in the Deegan-Packel index is inversely proportional to the number of voters that express support in ω. Hence, like the modified Johnston index, the probability model of the Deegan-Packel index implies that coalitions with more voters expressing support are less likely to occur.

[6] The author would like to point out that Holler doesn't advocate this as a realistic assumption, but acknowledges its usefulness in voting power calculations.

3.7 Summary

We've examined some of the most frequently used voting power calculation methods and described their criticality indicator functions. Let's recap on the different types of indicators that we've uncovered.

Increasing Criticality—This identifies situations in which the voter is able to change the outcome by increasing its support.

Decreasing Criticality—This identifies situations in which the voter is able to change the outcome by decreasing its support.

Total Criticality—This identifies situations in which the voter is able to change the outcome by either increasing, or decreasing its support.

So far, so good. However, we need to expand our notions of criticality a little further to handle those methods that restrict the initial or final way that a voter might vote.

Criticality 0—This criticality assumption requires that the voter must either start by initially voting "no", or that it must change its mind to end up ultimately voting "no".

Criticality δ—This criticality assumption places no restriction upon how the voter initially votes, or how it ultimately votes.

(Readers seeking a more in-depth discussion of criticality, and the motivation for the different types, are advised to consult Das (2011).)

Using these different notions of criticality, we can make affirmative statements with regards to what the standard voting power calculation techniques are actually calculating.

$$\begin{aligned}
\text{ShapleyShubik}_i &= \Pr(IC_i^0).\\
\text{Banzhaf}_i &= \Pr(TC_i^\delta).\\
\text{Straffin}_i &= \Pr(TC_i^\delta).\\
\text{Coleman Initiate Action}_i &= \Pr(IC_i^\delta \mid \text{Losing}).\\
\text{Coleman Prevent Action}_i &= \Pr(DC_i^\delta \mid \text{Winning}).\\
\text{Johnston}_i &= \Pr(DC_i^0).\\
\text{DeeganPackel}_i &= \Pr(DC_i^0).\\
\text{HollerPGI}_i &= \Pr(DC_i^0).
\end{aligned}$$

4 Simple Voting Games

In this section we concern ourselves with simple voting games, briefly put, we assume that all voters can only vote "yes" or "no", and that the outcome can only be Winning or Losing.

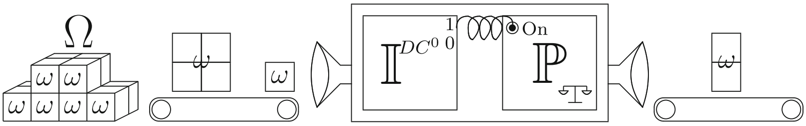


Fig. 4 Decreasing criticality 0

4.1 Decreasing Criticality **0**

Building upon the work we did counting coloured blocks, let's build a super-measurer to measure Decreasing Criticality 0 (Fig. 4).

If we were to turn our super-measurer on, and count different voting scenarios we would end up performing the following calculation.

$$\int_{\omega \in \Omega} \mathbb{I}^{DC_i^0}(\omega)\, \mathbb{P}(d\omega) = \Pr(DC_i^0).$$

Let's examine this machine in greater detail, we already know how the $\mathbb{P}$ machine works, so let's focus upon the $\mathbb{I}^{DC_i^0}$ machine. A Decreasing Criticality 0 event occurs when the game is currently winning, but then becomes losing when the voter i decreases its support. Therefore we need the following two conditions to hold true.

Condition 1—The current ω being measured must be winning.

Condition 2—A modified version of ω, which we will call ω' must be losing. ω' is the same as the ω, except that voter i has changed its mind, and is now voting "no".

Clearly, we need a way of identifying when an ω is winning. So let's make an indicator function for it.

$$\mathbb{I}^{Win}(\omega) = \begin{cases} 1 & \text{if } \omega \text{ is classified Winning;} \\ 0 & \text{otherwise.} \end{cases}$$

Using this new indicator function, we can now write down the Decreasing Criticality 0 indicator function.

$$\mathbb{I}^{DC_i^0}(\omega) = \mathbb{I}^{Win}(\omega) - \mathbb{I}^{Win}(\omega').$$

The proof of this is given by the following truth table.

$\mathbb{I}^{Win}(\omega)$	$\mathbb{I}^{Win}(\omega')$	$\mathbb{I}^{Win}(\omega) - \mathbb{I}^{Win}(\omega')$
0	0	0
0	1	N/A*
1	1	0
1	0	1

* It should be noted that the construction of ω' ensures that it is not possible for ω' to be winning, while ω is losing.

Let's make use of our new indicator for Decreasing Criticality 0 to write down what the super-measurer is measuring.

$$\Pr(DC_i^0) = \int_{\omega\in\Omega} \mathbb{I}^{DC_i^0}(\omega)\ \mathbb{P}(d\omega) = \int_{\omega\in\Omega} \mathbb{I}^{Win}(\omega) - \mathbb{I}^{Win}(\omega')\ \mathbb{P}(d\omega).$$

Just like any other normal integral, we can split this to give,

$$\Pr(DC_i^0) = \int_{\omega\in\Omega} \mathbb{I}^{Win}(\omega)\ \mathbb{P}(d\omega) - \int_{\omega\in\Omega} \mathbb{I}^{Win}(\omega')\ \mathbb{P}(d\omega).$$

Look at the first integral $\int_{\omega\in\Omega} \mathbb{I}^{Win}(\omega)\ \mathbb{P}(d\omega)$, this looks a lot like the expression we created for calculating the probability of a blue block,

$$\int_{\omega\in\Omega} \mathbb{I}^{Blue}(\omega)\ \mathbb{P}(d\omega) = \Pr(\text{Blue}).$$

By the same logic,

$$\int_{\omega\in\Omega} \mathbb{I}^{Win}(\omega)\ \mathbb{P}(d\omega) = \Pr(\text{Winning}).$$

Hence,

$$\Pr(DC_i^0) = \Pr(\text{Winning}) - \int_{\omega\in\Omega} \mathbb{I}^{Win}(\omega')\ \mathbb{P}(d\omega).$$

Now let's examine the second integral $\int_{\omega\in\Omega} \mathbb{I}^{Win}(\omega')\ \mathbb{P}(d\omega)$. This is simply the probability that ω' is winning. Hence,

$$\Pr(DC_i^0) = \Pr(\text{Winning}) - \Pr(\omega' \text{ is Winning}).$$

How do we interpret this? Recall that ω' is when voter i changes its mind to vote "no". As $\Pr(DC_i^0)$ is expressed as Pr(Winning) less $\Pr(\omega'$ is Winning), it is clearly the drop in likelihood of the game being winning when voter i changes to vote "no" in every situation.

So far so good, but we can simplify this expression further. Recall that the function $\mathbb{P}$ "weighs" a block, and $\int_{\Omega} \mathbb{P}(d\omega)$ is the process of "weighing" all the blocks. What if we took a block, and broke off a small piece, say we chipped off one corner? The small fragment we will call i, and the remainder of the block we will call $\omega^{N\setminus\{i\}}$. If you put both pieces onto the weighing scales, they would still weigh the same as the original block. Likewise, if you weighed the small i piece first, and the bigger $\omega^{N\backslash\{i\}}$ piece second, you could still calculate the weight of the original block by adding up the two results.

We can take this idea even further, we could break a corner off of every block $\omega \in \Omega$, and still calculate the total weight of Ω by first weighing all the small i pieces, and then weighing all the larger $\omega^{N\backslash\{i\}}$ pieces. Let's write down this process using integral notation.

$$\int_{\omega\in\Omega} \mathbb{I}^{Win}(\omega')\ \mathbb{P}(d\omega) = \int_{\omega^{N\backslash\{i\}}} \int_i \mathbb{I}^{Win}(\omega')\ \mu(di)\ \lambda(d\omega^{N\backslash\{i\}}).$$

In the above equation, we've replaced the $\mathbb{P}$ "weighing" machine with two new weighing machines; μ which specialises in weighing the small fragments, and λ which specialises in the larger fragments.[7] Therefore we can express $\Pr(DC_i^0)$ as,

$$\Pr(DC_i^0) = \Pr(\text{Winning}) - \int_{\omega^{N\backslash\{i\}}} \int_i \mathbb{I}^{Win}(\omega')\ \mu(di)\ \lambda(d\omega^{N\backslash\{i\}}).$$

There's one more change to make, recall that ω' is explicitly constructed as ω with voter i changing its vote to "no". Let's take ω', and break it into two fragments, i and $\omega^{N\backslash\{i\}}$, we can rename the i part to i_{no} (to show that i is always voting "no"). Doing this yields,

$$\Pr(DC_i^0) = \Pr(\text{Winning}) - \int_{\omega^{N\backslash\{i\}}} \int_i \mathbb{I}^{Win}(\omega^{N\backslash\{i\}} \times i_{no})\ \mu(di)\ \lambda(d\omega^{N\backslash\{i\}}).$$

Examine the inner integral over the variable i. The term inside, $\mathbb{I}^{Win}(\omega^{N\backslash\{i\}} \times i_{no})$, is constant with respect to i, so it can be brought outside of this integral to give,

$$\Pr(DC_i^0) = \Pr(\text{Winning}) - \int_{\omega^{N\backslash\{i\}}} \mathbb{I}^{Win}(\omega^{N\backslash\{i\}} \times i_{no}) \int_i \mu(di)\ \lambda(d\omega^{N\backslash\{i\}}).$$

Remember that our specialised "weighing" machines always add up to 1 if they count every element, i.e. $\int_i \mu(di) = 1$. Using this, we can remove it from our expression to give,

$$\Pr(DC_i^0) = \Pr(\text{Winning}) - \int_{\omega^{N\backslash\{i\}}} \mathbb{I}^{Win}(\omega^{N\backslash\{i\}} \times i_{no})\ \lambda(d\omega^{N\backslash\{i\}}).$$

[7] The new "weighing" machines have been simplified for the purposes of this example, they are actually sigma finite marginal measures, and are more correctly given by $\mu_{\omega^{N\backslash\{i\}}}(dx_i)\ \lambda(d\omega^{N\backslash\{i\}})$. (See Appendix 3.)

Let's briefly discuss $\int_{\omega^{N\setminus\{i\}}} \mathbb{I}^{Win}(\omega^{N\setminus\{i\}} \times i_{no})\ \lambda(d\omega^{N\setminus\{i\}})$. This is saying, take all the broken large fragments, and instead of putting back the original i you chipped off, replace it with an i_{no}, then calculate the probability of winning with this new "glued" together block. In other words, this is still Pr(ω' is winning).

Now let's examine the special weighing machine $\lambda(d\omega^{N\setminus\{i\}})$. This gives us the probability distribution of the other voters (i.e. everyone but voter i). If the other voters do not, or can not respond to how voter i voted, then their probability distribution would be unchanged if voter i voted "yes", or if it voted "no".[8] Therefore,

$$\lambda_{i_{yes}}(d\omega^{N\setminus\{i\}}) = \lambda_{i_{no}}(d\omega^{N\setminus\{i\}}) = \lambda(d\omega^{N\setminus\{i\}}).$$

And thus,

$$\Pr(DC_i^0) = \Pr(\text{Winning}) - \int_{\omega^{N\setminus\{i\}}} \mathbb{I}^{Win}(\omega^{N\setminus\{i\}} \times i_{no})\ \lambda_{i_{no}}(d\omega^{N\setminus\{i\}}).$$

Let's examine this last integral in some detail,

$$\int_{\omega^{N\setminus\{i\}}} \mathbb{I}^{Win}(\omega^{N\setminus\{i\}} \times i_{no})\ \lambda_{i_{no}}(d\omega^{N\setminus\{i\}}).$$

The "weighing" machine is called a sigma finite marginal measure. It assumes that voter i votes i_{no} and then calculates the probability distribution of the other voters. It has, in effect, marginalised out voter i. Now let's look at the indicator function, it is identifying when the game is winning given that voter i voted i_{no}. In essence, this integral is calculating the conditional probability of the game being winning, given that voter i voted "no". Ergo,

$$\int_{\omega^{N\setminus\{i\}}} \mathbb{I}^{Win}(\omega^{N\setminus\{i\}} \times i_{no})\ \lambda_{i_{no}}(d\omega^{N\setminus\{i\}}) = \Pr(\text{Winning} \mid i_{no}).$$

And hence,

$$\Pr(DC_i^0) = \Pr(\text{Winning}) - \Pr(\text{Winning} \mid i \text{ voted no}).$$

[8] It should be noted that the vast majority of real life voting games are structured to ensure that this is the case. For example, games where the votes are cast simultaneously, or games where they are cast anonymously. The key requirement is that the other voters cannot observe the actual event of voter i voting, and then react. We do not preclude scenarios in which voter i tells everyone how it intends to vote, providing the others do not actually see the vote taking place.

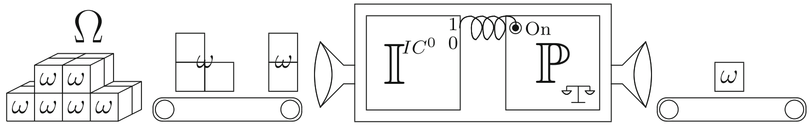


Fig. 5 Increasing criticality 0

4.1.1 Discussion

Before we move on we should take a moment to examine this result. Any voting power technique that calculates Decreasing Criticality 0, like Deegan-Packel, HollerPGI, and Johnston, is simply calculating the probability of the game being winning, less the probability of it being winning when voter i is set to voting "no". This is always true, irrespective of the underlying probability model assumed by the technique.

And for those voting games in which the probability distribution of the other voters is unaffected by the way voter i votes, this can be further simplified to the unconditional probability of the game being winning, less the conditional probability of the game being winning given that voter i has voted "no".

4.2 Increasing Criticality 0

Creating a super-measurer for Increasing Criticality 0 is a little more involved, so let's look at how this could be done (Fig. 5).

We already know how the $\mathbb{P}$ machine works, so let's focus upon the $\mathbb{I}^{IC_i^0}$ machine. By definition, for a little ω to be Increasing Criticality 0 we require three conditions to hold true.

Condition 1—As this is a Criticality 0 measure, the ω being measured must have voter i already expressing zero support (i.e. voting "no").

Condition 2—The ω being measured must be losing.

Condition 3—A modified version of ω, which we will call ω', must be winning. ω' is the same as ω, except that voter i has changed its mind, and is now voting "yes".

If we let i_{no} represent voter i voting "no", and we let the indicator function $\mathbb{I}^{i_{no}}$ be the indicator of voter i voting "no", then the indicator function for Increasing Criticality 0 is given by,

$$\mathbb{I}^{IC_i^0}(\omega) = \mathbb{I}^{i_{no}}(\omega)\left(\mathbb{I}^{Win}(\omega') - \mathbb{I}^{Win}(\omega)\right).$$

The proof of this is easily given by the following truth table.

$\mathbb{I}^{i_{no}}(\omega)$	$\mathbb{I}^{Win}(\omega')$	$\mathbb{I}^{Win}(\omega)$	$\mathbb{I}^{i_{no}}(\omega)\left(\mathbb{I}^{Win}(\omega') - \mathbb{I}^{Win}(\omega)\right)$
0	0	0	0
0	0	1	N/A*
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	N/A*
1	1	0	1
1	1	1	0

* The construction of ω' ensures that it is not possible for a ω' to be losing while ω is winning.

Integrating this indicator function with respect to the function $\mathbb{P}$ gives,

$$\Pr(IC_i^0) = \int_{\omega \in \Omega} \mathbb{I}^{i_{no}}(\omega) \times \left(\mathbb{I}^{Win}(\omega') - \mathbb{I}^{Win}(\omega)\right) \mathbb{P}(d\omega).$$

Let's multiply out the indicator functions, and split the resultant integral,

$$\Pr(IC_i^0) = \int_{\omega \in \Omega} \mathbb{I}^{i_{no}}(\omega)\, \mathbb{I}^{Win}(\omega')\, \mathbb{P}(d\omega) - \int_{\omega \in \Omega} \mathbb{I}^{i_{no}}(\omega)\, \mathbb{I}^{Win}(\omega)\, \mathbb{P}(d\omega).$$

When we multiply two indicator functions together we identify the intersection of their events. Hence,

$$\int_{\omega \in \Omega} \mathbb{I}^{i_{no}}(\omega)\, \mathbb{I}^{Win}(\omega')\, \mathbb{P}(d\omega) = \Pr((\omega' \text{ is Winning}) \cap i_{no}).$$

$$\int_{\omega \in \Omega} \mathbb{I}^{i_{no}}(\omega)\, \mathbb{I}^{Win}(\omega)\, \mathbb{P}(d\omega) = \Pr(\text{Winning} \cap i_{no}).$$

Therefore,

$$\Pr(IC_i^0) = \Pr((\omega' \text{ is Winning}) \cap i_{no}) - \Pr(\text{Winning} \cap i_{no}).$$

Let's consider what this means. We are focusing entirely upon those events in which the voter is voting "no". And we then calculate by how much the voter can increase the likelihood of the game being winning when the voter changes to vote "yes".

As we did with $\Pr(DC_i^0)$, we can express $\Pr(IC_i^0)$ in terms of easier to understand conditional probabilities, if we restrict ourselves to games in which the other voters are unaffected by the way voter i votes. Let's start again from the original indicator function.

$$\mathbb{I}^{IC_i^0}(\omega) = \mathbb{I}^{i_{no}}(\omega)\left(\mathbb{I}^{Win}(\omega') - \mathbb{I}^{Win}(\omega)\right).$$

This time instead of integrating with respect to $\mathbb{P}$ on the set Ω, let's integrate with respect to the sigma finite marginal measure $\lambda_{i_{no}}(d\omega^{N\setminus\{i\}})$, defined on the subsets of $\Omega^{N\setminus\{i\}}$, given voter i has voted i_{no}. This gives the conditional probability of IC_i^0, given that voter i has voted i_{no}.

$$\Pr(IC_i^0 \mid i_{no}) = \int_{\omega^{N\setminus\{i\}}} \mathbb{I}^{i_{no}}(\omega) \left(\mathbb{I}^{Win}(\omega') - \mathbb{I}^{Win}(\omega)\right) \lambda_{i_{no}}(d\omega^{N\setminus\{i\}}).$$

We know that the indicator function $\mathbb{I}^{i_{no}}(\omega)$ ensures that the only ω to be measured will have voter i voting i_{no}, allowing us to replace ω with $(\omega^{N\setminus\{i\}} \times i_{no})$, and we also know that ω' is $(\omega^{N\setminus\{i\}} \times i_{yes})$. Finally, we note that as we've already restricted ω to $(\omega^{N\setminus\{i\}} \times i_{no})$, and the integration is occurring over the sigma finite marginal measure $\lambda_{i_{no}}(d\omega^{N\setminus\{i\}})$, it follows that the indicator function $\mathbb{I}^{i_{no}}(\omega)$ has become redundant (because it will always show 1), and can be safely removed without loss.

$$\Pr(IC_i^0 \mid i_{no}) = \int_{\omega^{N\setminus\{i\}}} \mathbb{I}^{Win}(\omega^{N\setminus\{i\}} \times i_{yes}) - \mathbb{I}^{Win}(\omega^{N\setminus\{i\}} \times i_{no}) \lambda_{i_{no}}(d\omega^{N\setminus\{i\}}).$$

$$\Pr(IC_i^0 \mid i_{no}) = \int_{\omega^{N\setminus\{i\}}} \mathbb{I}^{Win}(\omega^{N\setminus\{i\}} \times i_{yes}) \lambda_{i_{no}}(d\omega^{N\setminus\{i\}}) - \int_{\omega^{N\setminus\{i\}}} \mathbb{I}^{Win}(\omega^{N\setminus\{i\}} \times i_{no}) \lambda_{i_{no}}(d\omega^{N\setminus\{i\}}).$$

In those games where the other voters are unable to react to the way voter i votes we have,

$$\lambda_{i_{yes}}(d\omega^{N\setminus\{i\}}) = \lambda_{i_{no}}(d\omega^{N\setminus\{i\}}).$$

Hence,

$$\Pr(IC_i^0 \mid i_{no}) = \int_{\omega^{N\setminus\{i\}}} \mathbb{I}^{Win}(\omega^{N\setminus\{i\}} \times i_{yes}) \lambda_{i_{yes}}(d\omega^{N\setminus\{i\}}) - \int_{\omega^{N\setminus\{i\}}} \mathbb{I}^{Win}(\omega^{N\setminus\{i\}} \times i_{no}) \lambda_{i_{no}}(d\omega^{N\setminus\{i\}}).$$

We interpret this to be,

$$\Pr(IC_i^0 \mid i_{no}) = \Pr(\text{Winning} \mid i_{yes}) - \Pr(\text{Winning} \mid i_{no}).$$

Next we multiply both sides by $\Pr(i_{no})$, which along with an application of Bayes' theorem gives,

$$\Pr(IC_i^0 \cap i_{no}) = \Pr(i_{no}) \times (\Pr(\text{Winning} \mid i_{yes}) - \Pr(\text{Winning} \mid i_{no})).$$

We know that, by definition, IC_i^0 must have voter i voting i_{no}, i.e. $IC_i^0 \cap i_{no} = IC_i^0$. Therefore,

$$\Pr(IC_i^0) = \Pr(i_{no}) \times (\Pr(\text{Winning} \mid i_{yes}) - \Pr(\text{Winning} \mid i_{no})).$$

4.2.1 Discussion

Any technique that calculates Increasing Criticality 0, such as the Shapley-Shubik index, only looks at those events in which voter i is voting "no". It calculates the probability of winning if it were to change to voting "yes", less the probability of winning if it stayed voting "no". This is true, irrespective of probability model.

And in those games where the other voters are unaffected by the way voter i votes, this can be expressed as the conditional probability of the game being winning given voter i votes "yes", less the conditional probability of the game being winning given voter i votes "no", all multiplied by the probability of the voter voting "no".

4.3 *Total Criticality* 0

As it is not possible to be both Increasing Criticality and Decreasing Criticality at the same time in the same ω, we know that they are mutually exclusive events. Hence, there is no need to create a new super-measurer for Total Criticality 0, we can simply add together the Increasing and Decreasing Criticality results.

$$\begin{aligned}\Pr(TC_i^0) = \Pr(\text{Winning}) - \Pr((\omega^{N\setminus\{i\}} \times i_{no}) \text{ is Winning}) +\\ \Pr(((\omega^{N\setminus\{i\}} \times i_{yes}) \text{ is Winning}) \cap i_{no}) - \Pr(\text{Winning} \cap i_{no}).\end{aligned}$$

And in those games where the other voters are unaffected by the way voter i votes, this can be simplified to,

$$\begin{aligned}\Pr(TC_i^0) = \Pr(\text{Winning}) - \Pr(\text{Winning} \mid i_{no}) +\\ \Pr(i_{no}) \times \left(\Pr(\text{Winning} \mid i_{yes}) - \Pr(\text{Winning} \mid i_{no})\right).\end{aligned}$$

4.4 *Increasing Criticality δ*

The indicator function for Increasing Criticality δ is given by,

$$\mathbb{I}^{IC_i^\delta}(\omega) = \mathbb{I}^{Win}(\omega^{N\setminus\{i\}} \times i_{yes}) - \mathbb{I}^{Win}(\omega).$$

The proof of this is given by the following truth table.

$\mathbb{I}^{Win}(\omega^{N\setminus\{i\}} \times i_{yes})$	$\mathbb{I}^{Win}(\omega)$	$\mathbb{I}^{Win}(\omega^{N\setminus\{i\}} \times i_{yes}) - \mathbb{I}^{Win}(\omega)$
0	0	0
0	1	N/A
1	1	0
1	0	1

Integrating this indicator function with respect to $\mathbb{P}$ yields,

$$\Pr(IC_i^{\delta}) = \Pr((\omega^{N\setminus\{i\}} \times i_{yes}) \text{ is Winning}) - \Pr(\text{Winning}).$$

This simplifies to the following, whenever the other voters are unaffected by the way voter i votes,

$$\Pr(IC_i^{\delta}) = \Pr(\text{Winning} \mid i_{yes}) - \Pr(\text{Winning}).$$

4.5 Decreasing Criticality δ

The indicator function for Decreasing Criticality δ is given by,

$$\mathbb{I}^{DC_i^{\delta}}(\omega) = \mathbb{I}^{Win}(\omega) - \mathbb{I}^{Win}(\omega^{N\setminus\{i\}} \times i_{no}).$$

The proof of this is given by the following truth table.

$\mathbb{I}^{Win}(\omega)$	$\mathbb{I}^{Win}(\omega^{N\setminus\{i\}} \times i_{no})$	$\mathbb{I}^{Win}(\omega) - \mathbb{I}^{Win}(\omega^{N\setminus\{i\}} \times i_{no})$
0	0	0
0	1	N/A
1	1	0
1	0	1

Integrating this indicator function with respect to $\mathbb{P}$ yields,

$$\Pr(DC_i^{\delta}) = \Pr(\text{Winning}) - \Pr((\omega^{N\setminus\{i\}} \times i_{no}) \text{ is Winning}).$$

This simplifies to the following, whenever the other voters are unaffected by the way voter i votes,

$$\Pr(DC_i^{\delta}) = \Pr(\text{Winning}) - \Pr(\text{Winning} \mid i_{no}).$$

4.6 Total Criticality δ

There is no need to create a new super-measurer for Total Criticality δ, because the expression for Total Criticality is simply the sum of both Increasing and Decreasing Criticality.

$$\Pr(TC_i^{\delta}) = \Pr((\omega^{N\setminus\{i\}} \times i_{yes}) \text{ is Winning}) - \Pr((\omega^{N\setminus\{i\}} \times i_{no}) \text{ is Winning}).$$

This simplifies to the following, whenever the other voters are unaffected by the way voter i votes,

$$\Pr(TC_i^{\delta}) = \Pr(\text{Winning} \mid i_{yes}) - \Pr(\text{Winning} \mid i_{no}).$$

Therefore, techniques like the Banzhaf measure, or the Straffin index, are simply calculating the conditional probability of the game being winning, given that voter i has voted "yes", less the conditional probability of the game being winning, given that voter i has voted "no".

4.7 Discussion

Let's recap what we've discovered so far.

$$\begin{aligned}
\Pr(IC_i^{0}) &= \Pr(((\omega^{N\setminus\{i\}} \times i_{yes}) \text{ is Winning}) \cap i_{no}) - \Pr(\text{Winning} \cap i_{no}).\\
\Pr(DC_i^{0}) &= \Pr(\text{Winning}) - \Pr((\omega^{N\setminus\{i\}} \times i_{no}) \text{ is Winning}).\\
\Pr(TC_i^{0}) &= \Pr(\text{Winning}) - \Pr((\omega^{N\setminus\{i\}} \times i_{no}) \text{ is Winning}) +\\
&\quad \Pr(((\omega^{N\setminus\{i\}} \times i_{yes}) \text{ is Winning}) \cap i_{no}) - \Pr(\text{Winning} \cap i_{no}).\\
\Pr(IC_i^{\delta}) &= \Pr((\omega^{N\setminus\{i\}} \times i_{yes}) \text{ is Winning}) - \Pr(\text{Winning}).\\
\Pr(DC_i^{\delta}) &= \Pr(\text{Winning}) - \Pr((\omega^{N\setminus\{i\}} \times i_{no}) \text{ is Winning}).\\
\Pr(TC_i^{\delta}) &= \Pr((\omega^{N\setminus\{i\}} \times i_{yes}) \text{ is Winning}) - \Pr((\omega^{N\setminus\{i\}} \times i_{no}) \text{ is Winning}).
\end{aligned}$$

And in those voting games where the other voters are unaffected by the way voter i votes (for example, in games where the votes are cast simultaneously or anonymously),

$$\begin{aligned}
\Pr(IC_i^{0}) &= \Pr(i_{no}) \times (\Pr(\text{Winning} \mid i_{yes}) - \Pr(\text{Winning} \mid i_{no})).\\
\Pr(DC_i^{0}) &= \Pr(\text{Winning}) - \Pr(\text{Winning} \mid i_{no}).\\
\Pr(TC_i^{0}) &= \Pr(\text{Winning}) - \Pr(\text{Winning} \mid i_{no}) +
\end{aligned}$$

$$\begin{aligned}
&\Pr(i_{no}) \times \left(\Pr(\text{Winning} \mid i_{yes}) - \Pr(\text{Winning} \mid i_{no})\right).\\
\Pr(IC_i^{\delta}) &= \Pr(\text{Winning} \mid i_{yes}) - \Pr(\text{Winning}).\\
\Pr(DC_i^{\delta}) &= \Pr(\text{Winning}) - \Pr(\text{Winning} \mid i_{no}).\\
\Pr(TC_i^{\delta}) &= \Pr(\text{Winning} \mid i_{yes}) - \Pr(\text{Winning} \mid i_{no}).
\end{aligned}$$

Methods that calculate Decreasing Criticality (both 0 and δ), like the Deegan-Packel, Johnston, and HollerPGI indices, are simply calculating the probability of the game being winning, less the probability of it being winning when voter i votes "no". Or to put it another way, they are measuring the complete ability of the voter to prevent an outcome they don't want. If we were looking for a way of measuring the importance of a voter to a winning coalition, then we couldn't have dreamed up a more intuitive method than this.

Methods that calculate Increasing Criticality 0, like Shapley-Shubik, focus exclusively upon those events in which the voter is voting "no". This proportionality to $\Pr(i_{no})$ presents a challenge to these methods. To understand why this is such a problem, think about a game where the voter can abstain in addition to voting "yes" or "no". Arguably, if a voter does not have any inherent bias, their probability of voting "no" should now tend towards $\frac{1}{3}$. Now, let's imagine a game where the voter can abstain, or vote "yes", "no", and "maybe".[9] In this scenario the probability of voting "no" should tend towards $\frac{1}{4}$. In the most general case, where a voter expresses their vote by selecting from a continuous range of options (for example, if they had to rate their approval for a motion with a percentage), then $\Pr(i_{no}) \to 0$, and accordingly $\Pr(IC_i^0) \to 0$. Giving rise to the distinct possibility that such a technique would suggest that even dictators have zero voting power.[10]

Finally, let's discuss the Total Criticality δ methods like Banzhaf and Straffin. If we believe that voting power is the ability of a voter to influence the outcome of a vote, then the total (maximum) influence a voter could exert is given by the probability of that outcome given the voter tries its hardest to make the outcome more likely, less the probability of that outcome given the voter tries its hardest to make the outcome less likely. Which is precisely what these methods are calculating.

It is comforting to realise that the many different standard techniques have been calculating these common sense probabilities all along. Nothing could have been worse than finding out that they were calculating something strange, and nonsensical. Fortunately, this isn't the case. And all the investment made, and time spent, analysing voting power using the standard techniques has not been wasted. In fact, the techniques have been doing exactly what we hoped for, and now we can prove it.

[9] Abstention, is not the same as "maybe", see Das (2008) for details.

[10] This could only be avoided with the use of a biased probability distribution which imposed a disproportionately high likelihood of the voter voting "no".

5 Conclusion

Using a simple block counting example, this paper showed how to construct a measuring machine to calculate any statistic. Adapting the machine for voting power enabled us to calculate any voting power measure, irrespective of the underlying probability model. In the process we were able to establish exactly what the different voting power techniques were calculating.

We discovered that the Banzhaf measure is calculating,

$$\Pr(\text{Winning} \mid i \text{ votes yes}) - \Pr(\text{Winning} \mid i \text{ votes no}).$$

That the Shapley-Shubik index is calculating,

$$\Pr(i \text{ votes no}) \times (\Pr(\text{Winning} \mid i \text{ votes yes}) - \Pr(\text{Winning} \mid i \text{ votes no})) .$$

And that the two Coleman indices are calculating,

$$\frac{\Pr(\text{Winning} \mid i \text{ votes yes}) - \Pr(\text{Winning})}{1 - \Pr(\text{Winning})}$$
$$\frac{\Pr(\text{Winning}) - \Pr(\text{Winning} \mid i \text{ votes no})}{\Pr(\text{Winning})}.$$

In fact, using the block counting methodology, we saw how to construct any of the commonly applied voting power techniques using just four different probabilities:

$$\Pr(i \text{ votes no}), \quad \Pr(\text{Winning}),$$

$$\Pr(\text{Winning} \mid i \text{ votes yes}) \text{ and } \Pr(\text{Winning} \mid i \text{ votes no}).$$

The nice thing about this representation is that it allows us to decouple the voting power measure from the probability model of the game. Why is this so important? Not only does this allow us to analyse voting games that might have non-standard probability models, but it also allows us to use different, and potentially more accurate, probability models.[11] Which, in turn, can generate more accurate voting power statistics.

The results given in the main body of the paper are generalised within the Appendix to encompass games with multiple voter choices, multiple possible game outcomes, and arbitrary decision rules. Excluding some minor notational change, the results have stayed the same. Hence, we now have all the necessary theoretical

[11] The development of more realistic probability models will no doubt become a huge challenge for the future.

tools to analyse any voting game we desire, and, crucially, we now know exactly what these tools are calculating.

We started this paper by repeating the publicly stated desire of Machover for voting power to adopt a more probabilistic approach, and the expressed fear of Felsenthal that he would not see an explanation of what voting power was actually calculating within his lifetime. We end this paper with the modest hope that their desires have been fulfilled, and their fears allayed.

Appendix 1: Abstentions and More

The main body of the paper examined voting games in which the voters were only allowed to vote "yes" or "no". The concept of abstention was completely ignored. In this section, not only will we incorporate abstentions into our block counting methodology, but we will also expand the number of voting choices available to the voter.

Abstention

So let's start with the most basic change, instead of allowing a voter to vote "yes" or "no", we will now allow "yes", "no", or "abstain". The process of creating voting power measures is the same as we saw previously, first we will construct our indicator functions, and then integrate them using $\mathbb{P}$.

Indicator Functions

The indicator functions for the different criticalities are given below. We skip the tiresome listing of truth tables and simply state the functions instead.

$$\mathbb{I}^{DC_i^0}(\omega) = \mathbb{I}^{Win}(\omega) - \mathbb{I}^{Win}(\omega^{N\backslash\{i\}} \times i_{no}).$$

$$\mathbb{I}^{IC_i^0}(\omega) = \mathbb{I}^{i_{no}}(\omega)\left(\mathbb{I}^{Win}(\omega^{N\backslash\{i\}} \times i_{yes}) - \mathbb{I}^{Win}(\omega)\right).$$

$$\mathbb{I}^{DC_i^\delta}(\omega) = \mathbb{I}^{Win}(\omega) - \mathbb{I}^{Win}(\omega^{N\backslash\{i\}} \times i_{no}).$$

$$\mathbb{I}^{IC_i^\delta}(\omega) = \mathbb{I}^{Win}(\omega^{N\backslash\{i\}} \times i_{yes}) - \mathbb{I}^{Win}(\omega).$$

Remarkably, these are exactly the same indicator functions we used in the simple "yes/no" voting games. The addition of abstention hasn't changed the indicator functions.

$\mathbb{P}$ and Ω

We can create the new set Ω relatively easily. We simply take every voter in the game and then "combine" them together to create every possible combination of voting actions. What do we mean by this? Imagine a game with two voters, Voter 1 can vote "yes" or "no", and Voter 2 can vote 'yes", "no", and "abstain". If we were to "combine" them together we would end up with the following list of possible voter actions,

Voter 1	Voter 2
"yes"	"yes"
"yes"	"abstain"
"yes"	"no"
"no"	"yes"
"no"	"abstain"
"no"	"no"

(Each element of the set $\omega \in \Omega$ is represented as a separate line in this table.)

This example shows how to modify Ω (and by extension $\mathbb{P}$) to incorporate new voting choices. All we need do is "combine" every possible voter choice with every other possible voter choice, to create an enlarged set Ω.

Integrating the Indicators

If there is one thing we've learnt from block counting, it's that, even if Ω and $\mathbb{P}$ change, providing the indicator is unchanged, the statistic being calculated must be unchanged. When we added abstentions we didn't need to change the indicators, so it follows that, even in a game with abstentions, the voting power measures are given by the expressions in Sect. 4.7.

And More…

OK, so if we added abstentions so easily into our methodology perhaps we could do more? What if we allow extra voting options like "25 % in favour", or "maybe"? Why not take this idea to its logical conclusion and let the voters select from a possibly infinite range of options?

The Indicator Functions

Giving a voter an infinite range of options to choose from means that we might no longer have an option that we can definitively call "yes", or an option that we can definitively call "no". So instead we define two new options called i_{max} and i_{min}. These are the generalised equivalents of voting "yes" and "no", and change the likelihood of voter i's desired outcome by the greatest amount. Using this new terminology we can give the new indicator functions as,

$$\begin{aligned}
\mathbb{I}^{DC_i^0}(\omega) &= \mathbb{I}^{Win}(\omega) - \mathbb{I}^{Win}(\omega^{N\setminus\{i\}} \times i_{min}). \\
\mathbb{I}^{IC_i^0}(\omega) &= \mathbb{I}^{i_{min}}(\omega)\left(\mathbb{I}^{Win}(\omega^{N\setminus\{i\}} \times i_{max}) - \mathbb{I}^{Win}(\omega)\right). \\
\mathbb{I}^{DC_i^\delta}(\omega) &= \mathbb{I}^{Win}(\omega) - \mathbb{I}^{Win}(\omega^{N\setminus\{i\}} \times i_{min}). \\
\mathbb{I}^{IC_i^\delta}(\omega) &= \mathbb{I}^{Win}(\omega^{N\setminus\{i\}} \times i_{max}) - \mathbb{I}^{Win}(\omega).
\end{aligned}$$

$\mathbb{P}$ and Ω

Adding potentially infinite options to each voter clearly changes Ω and $\mathbb{P}$. Just like before, all we do is "combine" the different voters to create the new set Ω, and the new $\mathbb{P}$.

Integrating the New Indicators

Once again, the new $\mathbb{P}$ functions will not affect the statistic being calculated, however we are using slightly different indicators. When we integrate these new indicators we get the following,

$$\begin{aligned}
\Pr(IC_i^0) &= \Pr(((\omega^{N\setminus\{i\}} \times i_{max}) \text{ is Winning}) \cap i_{min}) - \Pr(\text{Winning} \cap i_{min}). \\
\Pr(DC_i^0) &= \Pr(\text{Winning}) - \Pr((\omega^{N\setminus\{i\}} \times i_{min}) \text{ is Winning}). \\
\Pr(TC_i^0) &= \Pr(\text{Winning}) - \Pr((\omega^{N\setminus\{i\}} \times i_{min}) \text{ is Winning}) + \\
&\quad \Pr(((\omega^{N\setminus\{i\}} \times i_{max}) \text{ is Winning}) \cap i_{min}) - \Pr(\text{Winning} \cap i_{min}). \\
\Pr(IC_i^\delta) &= \Pr((\omega^{N\setminus\{i\}} \times i_{max}) \text{ is Winning}) - \Pr(\text{Winning}). \\
\Pr(DC_i^\delta) &= \Pr(\text{Winning}) - \Pr((\omega^{N\setminus\{i\}} \times i_{min}) \text{ is Winning}). \\
\Pr(TC_i^\delta) &= \Pr((\omega^{N\setminus\{i\}} \times i_{max}) \text{ is Winning}) - \Pr((\omega^{N\setminus\{i\}} \times i_{min}) \text{ is Winning}).
\end{aligned}$$

And in those voting games where other voters are not affected by the way voter i votes,

$$\begin{aligned}
\Pr(IC_i^0) &= \Pr(i_{min}) \times (\Pr(\text{Winning} \mid i_{max}) - \Pr(\text{Winning} \mid i_{min})). \\
\Pr(DC_i^0) &= \Pr(\text{Winning}) - \Pr(\text{Winning} \mid i_{min}).
\end{aligned}$$

$$
\begin{aligned}
\Pr(TC_i^0) &= \Pr(\text{Winning}) - \Pr(\text{Winning} \mid i_{min}) + \\
&\quad \Pr(i_{min}) \times (\Pr(\text{Winning} \mid i_{max}) - \Pr(\text{Winning} \mid i_{min}))\,. \\
\Pr(IC_i^\delta) &= \Pr(\text{Winning} \mid i_{max}) - \Pr(\text{Winning}). \\
\Pr(DC_i^\delta) &= \Pr(\text{Winning}) - \Pr(\text{Winning} \mid i_{min}). \\
\Pr(TC_i^\delta) &= \Pr(\text{Winning} \mid i_{max}) - \Pr(\text{Winning} \mid i_{min}).
\end{aligned}
$$

The keen eyed reader will no doubt have spotted that these expressions are the same as the ones we generated for the simple "yes/no" voting games with the terms i_{no} and i_{yes} replaced by i_{min} and i_{max}.

Appendix 2: Multiple Outcomes and Complex Non-monotonic Decision Rules

Up to now we have been dealing with games that can be either "Winning" or "Losing". But can we generalise our ideas to encompass more complex games? Games with more than two outcomes? Perhaps games that give some kind of ranking of alternatives? Once again, we find that we can do this, and more, with a minimum of fuss. But before we look at expanding the number of possible outcomes, let's discuss the voting decision rule. Even though it was never explicitly stated before, there is no restriction on the decision rule. There is no requirement for it to be weighted, monotonic, or in any way sensible. It could be the most complex, non-monotonic rule you can think of. It will not affect our block counting methodology.

Now, back to expanding the number of possible outcomes. If we have a game with more than two mutually exclusive outcomes, then it becomes necessary to stipulate with respect to which particular outcome power is being measured. The reason for this is simple, in the most general types of games, with the potential for arbitrarily complex decision rules, the power of a voter might change from outcome to outcome.

Actually, this doesn't complicate things very much. All we need to do is change our indicator functions slightly. We now have to specify with respect to which particular outcome we are measuring criticality. We will use the symbol O to represent the specified outcome. And we also need to change the definitions of voter actions i_{min} and i_{max} so that they are given with respect to outcome O. We will use the symbols i_{min}^O and i_{max}^O to do this.

$$
\begin{aligned}
\mathbb{I}^{O\_DC_i^0}(\omega) &= \mathbb{I}^O(\omega) - \mathbb{I}^O(\omega^{N\setminus\{i\}} \times i_{min}^O). \\
\mathbb{I}^{O-IC_i^0}(\omega) &= \mathbb{I}^{i_{min}^O}(\omega)\left(\mathbb{I}^O(\omega^{N\setminus\{i\}} \times i_{max}^O) - \mathbb{I}^O(\omega)\right). \\
\mathbb{I}^{O\_DC_i^\delta}(\omega) &= \mathbb{I}^O(\omega) - \mathbb{I}^O(\omega^{N\setminus\{i\}} \times i_{min}^O). \\
\mathbb{I}^{O\_IC_i^\delta}(\omega) &= \mathbb{I}^O(\omega^{N\setminus\{i\}} \times i_{max}^O) - \mathbb{I}^O(\omega).
\end{aligned}
$$

Integrating these new indicators gives us the following,

$$\begin{aligned}
\Pr(O_IC_i^0) &= \Pr(((\omega^{N\setminus\{i\}} \times i_{max}^O) \text{ is } O) \cap i_{min}^O) - \Pr(O \cap i_{min}^O).\\
\Pr(O_DC_i^0) &= \Pr(O) - \Pr((\omega^{N\setminus\{i\}} \times i_{min}^O) \text{ is } O).\\
\Pr(O_TC_i^0) &= \Pr(O) - \Pr((\omega^{N\setminus\{i\}} \times i_{min}^O) \text{ is } O) +\\
&\quad \Pr(((\omega^{N\setminus\{i\}} \times i_{max}^O) \text{ is } O) \cap i_{min}^O) - \Pr(O \cap i_{min}^O).\\
\Pr(O_IC_i^\delta) &= \Pr((\omega^{N\setminus\{i\}} \times i_{max}^O) \text{ is } O) - \Pr(O).\\
\Pr(O_DC_i^\delta) &= \Pr(O) - \Pr((\omega^{N\setminus\{i\}} \times i_{min}^O) \text{ is } O).\\
\Pr(O_TC_i^\delta) &= \Pr((\omega^{N\setminus\{i\}} \times i_{max}^O) \text{ is } O) - \Pr((\omega^{N\setminus\{i\}} \times i_{min}^O) \text{ is } O).
\end{aligned}$$

And in those voting games where other voters are not affected by the way voter i votes,

$$\begin{aligned}
\Pr(O_IC_i^0) &= \Pr(i_{min}^O) \times (\Pr(O \mid i_{max}^O) - \Pr(O \mid i_{min}^O)).\\
\Pr(O_DC_i^0) &= \Pr(O) - \Pr(O \mid i_{min}^O).\\
\Pr(O_TC_i^0) &= \Pr(O) - \Pr(O \mid i_{min}^O) +\\
&\quad \Pr(i_{min}^O) \times \left(\Pr(O \mid i_{max}^O) - \Pr(O \mid i_{min}^O)\right).\\
\Pr(O_IC_i^\delta) &= \Pr(O \mid i_{max}^O) - \Pr(O).\\
\Pr(O_DC_i^\delta) &= \Pr(O) - \Pr(O \mid i_{min}^O).\\
\Pr(O_TC_i^\delta) &= \Pr(O \mid i_{max}^O) - \Pr(O \mid i_{min}^O).
\end{aligned}$$

Once again we see that these are almost the same expressions we generated for the simple "yes/no" voting games. The most obvious difference being that power is now specified with respect to a given outcome O, and the idea of voting "yes" or "no" has been replaced with the action that most favours outcome O, and the action that least favours outcome O.

With these expressions we can now calculate voting power statistics in practically any voting game we desire. The decision rule of the game can be arbitrarily complex, the game can have many different possible outcomes, and every voter can have an infinite range of different voting actions to choose from.

Appendix 3: Definitions

This paper has deliberately simplified some of the more rigorous mathematical terms in order to ease comprehension of the material. In this section we give the required formal definitions.

Rather than restrict our analysis to a specific voting system, we will introduce here the concept of a generalised voting game. This generalised voting game encompasses all possible voting games of interest, in that it allows for any voting rule, any number of possible voting outcomes, and any probability distribution of the voters. In keeping with the spirit of generalisation, we will henceforth refer to the voters as players. The definitions of probability and product spaces are taken from Pollard (2003).

Definition 1. A **player** is a probability space $(\mathcal{X}_i, \mathcal{A}_i, \mathbb{P}_i)$, where $\mathcal{X}_i$ is a set, $\mathcal{A}_i$ is a sigma-field of subsets of $\mathcal{X}_i$, and $\mathbb{P}_i$ is a countably additive, nonnegative measure with $\mathbb{P}_i(\mathcal{X}_i) = 1$. Given a set of N players, where $|N| = n$, the set of all ordered n-tuples $(x_1, \ldots, x_n)$, with $x_j \in \mathcal{X}_j$ for each $j \in 1, \ldots n$ is denoted as $\mathcal{X}_1 \times \cdots \times \mathcal{X}_n$ and abbreviated to Ω^N. Given a player i, the set of all ordered $(n-1)$-tuples $(x_1, \ldots, x_{i-1}, x_{i+1}, x_n)$, with $x_j \in \mathcal{X}_j$ for each $j \in 1, \ldots, i-1, i+1, \ldots n$ is denoted as $\mathcal{X}_1 \times \cdots \times \mathcal{X}_{i-1} \times \mathcal{X}_{i+1} \times \cdots \times \mathcal{X}_n$ and abbreviated to $\Omega^{N\setminus\{i\}}$. The action of creating a single $(n-1)$-tuple, denoted as $\omega^{N\setminus\{i\}}$, from a single n-tuple ω^N by removing the element x_i is represented as $\omega^N \setminus x_i$. The action of creating a single n-tuple, denoted as ω^N, from a single $(n-1)$-tuple $\omega^{N\setminus\{i\}}$ by adding an element $x_i \in \mathcal{X}_i$ is represented as $\omega^{N\setminus\{i\}} \times x_i$.

Definition 2. Given a set of N players, where $|N| = n$, a set of the form $A_1 \times \cdots \times A_n = \{(x_1, \ldots, x_n) \in \mathcal{X}_1 \times \cdots \times \mathcal{X}_n : x_i \in A_i \text{ for each } i\}$, with $A_i \in \mathcal{A}_i$ for each i, is called a measurable rectangle. The product sigma field $\mathcal{A}_1 \times \cdots \times \mathcal{A}_n$ on $\mathcal{X}_1 \times \cdots \times \mathcal{X}_n$ is defined to be the sigma field generated by all measurable rectangles. Let the product space $(\mathcal{X}_1 \times \cdots \times \mathcal{X}_n, \mathcal{A}_1 \times \cdots \times \mathcal{A}_n)$ be denoted as $(\Omega, \mathcal{F})$.

Definition 3. A **generalised voting game** is a quadruple $(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{W})$ such that $(\Omega, \mathcal{F}, \mathbb{P})$ is the product space generated by a set of N players, $\mathbb{P}$ is the product measure, and $\mathcal{W}$ is a $\mathcal{F} \setminus \mathcal{O}$ measurable function, where the elements $O \in \mathcal{O}$ are called **outcomes**. Such a game is denoted as a **GVG**$(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{W})$.

Definition 4. For a $GVG(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{W})$, a player i is **increasingly critical** with respect to an outcome $O \in \mathcal{O}$ in an event $\omega^N \in \Omega^N$ if, and only if, $\mathcal{W}(\omega^N) \neq O$ and there exists an $\{x_i'\} \in \mathcal{X}_i$ such that $\mathcal{W}((\omega^N \setminus \{x_i\}) \times \{x_i'\}) = O$. Let O_IC_i denote the set of increasingly critical events for a player i with respect to an outcome O.

Definition 5. For a $GVG(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{W})$, a player i is **decreasingly critical** with respect to an outcome $O \in \mathcal{O}$ in an event $\omega^N \in \Omega^N$ if, and only if, $\mathcal{W}(\omega^N) = O$ and there exists an $\{x_i'\} \in \mathcal{X}_i$ such that $\mathcal{W}((\omega^N \setminus \{x_i\}) \times \{x_i'\}) \neq O$. Let O_DC_i denote the set of decreasingly critical events for a player i with respect to an outcome O.

Definition 6. For a $GVG(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{W})$, a player i is **totally critical** with respect to an outcome $O \in \mathcal{O}$ in an event $\omega^N \in \Omega^N$ if it is either increasingly critical or decreasingly critical, with respect to the aforementioned outcome and event. Let O_TC_i denote the set of totally critical events for a player i with respect to an outcome O. For any given event ω^N, it is not possible to be simultaneously both increasingly and decreasingly critical with respect to a given outcome O, therefore $(O\_IC_i \cap O\_DC_i) = \emptyset$.

Definition 7. Criticality δ—With this assumption there is no restriction on how player i can vote between the two different events that define it as critical. The set of criticality δ increasingly critical events for player i, with respect to an outcome O, is denoted by $O\_IC_i^{\delta}$, and the set of criticality δ decreasingly critical events for player i, with respect to an outcome O, is denoted by $O\_DC_i^{\delta}$.

Definition 8. Criticality 0—With this assumption one of the two events that define player i as being critical must have player i voting with its lowest possible support for outcome O. The set of criticality 0 increasingly critical events for player i, with respect to an outcome O, is denoted by $O\_IC_i^{0}$, and the set of criticality 0 decreasingly critical events for player i, with respect to an outcome O, is denoted by $O\_DC_i^{0}$.

In simple "yes/no" voting games Criticality 0 and Criticality δ are equivalent. However, if any of the players are allowed to abstain this equivalence will be lost, and it will be necessary to understand which criticality assumption you wish to measure.

Definition 9. For a $GVG(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{W})$, a player i, and an outcome $O \in \mathcal{O}$, let $\mathbb{I}^O : \Omega^N \rightarrow \{\{0\}, \{1\}\}$ be the indicator function that an event ω^N is classified as outcome O, i.e. when $\mathcal{W}(\omega^N) = O$. Then, given an $\omega^{N\setminus\{i\}} \in \Omega^{N\setminus\{i\}}$, define $\{x_i^{O_{\max}}\}$ such that for all $x_i \in \mathcal{X}_i$,

$$\mathbb{I}^O\left(\omega^{N\setminus\{i\}} \times \{x_i^{O_{\max}}\}\right) \geq \mathbb{I}^O\left(\omega^{N\setminus\{i\}} \times x_i\right).$$

Likewise, define $\{x_i^{O_{\min}}\}$ such that for all $x_i \in \mathcal{X}_i$,

$$\mathbb{I}^O\left(\omega^{N\setminus\{i\}} \times \{x_i^{O_{\min}}\}\right) \leq \mathbb{I}^O\left(\omega^{N\setminus\{i\}} \times x_i\right).$$

$x_i^{O_{\min}}$ and $x_i^{O_{\max}}$ are generalised equivalents of voting "yes" and "no". They need not be unique elements within $\mathcal{X}_i$, and could instead be subsets. Should this turn out to be the case, the elements $\{x_i^{O_{\min}}\}$ and $\{x_i^{O_{\max}}\}$ can be taken as any appropriate element within said subsets.

References

Banzhaf, J. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review, 19,* 317–342.

Coleman, J. S. (1971). Control of collectives and the power of a collectivity to act. In B. Lieberman (Ed.), *Social choice* (pp. 192–225). New York: Gordon and Breach.

Das, S. (2008). *Class conditional voting probabilities*. Ph.D. thesis, Birkbeck College, University of London, London.

Das, S. (2011). Criticality in games with multiple levels of approval. *Social Choice and Welfare, 37,* 373–395. doi:10.1007/s00355-010-0488-z

Deegan, J., & Packel, E. (1978). A new index of power for simple n-person games. *International Journal of Game Theory, 7,* 113–123.

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

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

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

Pollard, D. (2003). *A user's guide to measure theoretic probability*. Cambridge: Cambridge University Press.

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

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

Straffin, P. (1978). Probability models for power indices. In P. C. Ordeshook (Ed.), *Game theory and political science* (pp. 477–510). New York: New York University Press.

Voting Power and Probability

Claus Beisbart

1 Introduction

One main aim of the seminal book "The Measurement of Voting Power: Theory and Practice, Problems and Paradoxes" (Felsenthal and Machover 1998) is to quantify voting power. The voting power of a citizen is defined as

> *the extent to which the member is able to control the outcome of a division of the board.* (Felsenthal and Machover 1998, p. 35).

Voting power is assumed to be normatively significant, for instance because voting systems should afford the same share of voting power to each citizen.[1]

To motivate their preferred measure of voting power, Felsenthal and Machover (1998) write (p. 36):

> How can this idea [the idea of voting power] be explicated mathematically? An obvious way – arguably the *only* reasonable way – of doing so is in terms of probability: the voting power of voter *a* can be formally defined as the probability of *a* being in a position to affect the outcome of a division.

[1] See e.g. (Felsenthal and Machover 2000, p. 17). In this paper, I will focus on I-power, which is about influence, while I bracket P-power, which is about prizes. See (Felsenthal and Machover 1998, p. 36) for the definition of both types of power.

C. Beisbart (✉)
Institut für Philosophie, Universität Bern, Länggassstr. 49a, CH-3012 Bern, Switzerland
e-mail: Claus.Beisbart@philo.unibe.ch

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare, DOI 10.1007/978-3-319-05158-1_6,

The crucial proposal of this passage is that the voting power of a voter be quantified using a *probability*, viz. the probability that the voter is pivotal. This very probability is usually calculated from a probability model over all possible voting profiles or coalitions. Felsenthal and Machover end up with defending the *Penrose/Banzhaf* measure of voting power, i.e., they quantify voting power with the probability that a voter is pivotal under the so-called *Bernoulli model* over voting profiles. The Bernoulli model assigns the same probability to each possible coalition.[2]

The aim of this paper is to discuss the transition from the definition of voting power (first quotation) to the idea that voting power is measured using a probability (second quotation). My question is as follows.

What kind of probability is the probability that measures voting power?

Put differently, the question is what "probability" means in the measurement of voting power. Is the probability a degree of belief or a real-world chance or something else? Twentieth-century philosophy has featured a rich discussion about the very notion of probability, and various analyses or interpretations of probability have been suggested. While some philosophers prefer objectivist accounts of probabilities, under which probabilities are e.g. relative frequencies or propensities, other authors argue for a subjectivist understanding of probabilities, or at least of some of them. My paper will draw on the general philosophical discussion about probabilities to deepen our understanding of voting power.[3]

Some people may object that an uncontroversial definition of probabilities is available because mathematicians define the notion of probability using the axioms of the probability calculus.[4] However, the mathematical definition of probability does not suffice for the purposes of this paper because it lacks any non-mathematical significance. When we equalize the voting powers of the citizens, we do not want to equalize arbitrary numbers that obey the axioms of the probability calculus. Rather, we are interested in a normatively significant feature of a voting system. In a similar way, scientists who put a probability of 1/3 on the outcome of an experiment take this probability to have significance that stretches beyond the realm of the purely mathematical. Philosophers of science examine what exactly this probability is. Likewise, I wish to examine what exactly the probability of pivotality is in voting theory.

Although my main aim is one of clarification, the analysis of this paper offers more than a purely philosophical exercise. The analysis has immediate consequences for the way voting power should be calculated in practice. As is

---

[2] See Felsenthal and Machover (1998), Definition 3.1.1 on p. 37 for this model.

[3] There is no presumption that all probabilities should be understood in the same way; rather, each probabilistic statement calls for its own interpretation. My focus here is entirely on probabilities that arise in the measurement of voting power. See (Hájek 1997, pp. 210–211) and (Hájek 2010, Sect. 2) for useful methodological remarks about the interpretation of probabilities. See Gillies (2000) for a text book on the philosophy of probability and Eagle (2011) for an anthology.

[4] See Kolmogorov (1956) for a famous version of the axioms.

well-known, the axioms of the probability calculus do not uniquely fix the values of the probabilities of most events.[5] Additional considerations are needed to set the values of probabilities. What these considerations are depends on the meaning of "probability" in the case at hand. For instance, if a probability is a relative frequency, then the value of the probability has to be identified with that of a relative frequency. To set the value of a probability we have thus to understand what the probability is.

As it happens, the question of how the values of probabilities should be set is at the center of two foundational issues in the literature about voting power. The *first issue* is whether voting power should be quantified a priori, i.e., without using empirical data, or a posteriori. The first option leads to aprioristic measures of voting power, which do not draw on empirical data, such as the Penrose measure. But recently, a number of authors have been attracted by the idea that voting power may be quantified using empirical data. For instance, Beisbart and Bovens (2008) calculate the probability of pivotality using a model that they have fitted to data from past elections in the voting system under consideration. Machover (2007) challenges this approach; Kaniovski and Leech (2009), Beisbart (2010) and Bovens and Beisbart (2011) make different suggestions to quantify power on the basis of empirical data. Despite this interest in a posteriori measures of voting power, some authors have argued that only a priori measures of voting power are relevant when we take a normative stance on voting systems (Felsenthal and Machover 1998, pp. 37–38; Laruelle and Valenciano 2005). To decide between a priori and a posteriori measures of voting power, we should know what type of probabilities the measures of voting power are. Such a decision has of course to be sensitive to the uses of voting power measures because different uses may require different interpretations of the probabilities.

The *second issue* only concerns a priori measures of voting power. Even if we agree that the values of the probability of pivotality should be calculated without any use of empirical information, there are several ways to do so. There are at least two probabilistic models over the voting profiles that are used to quantify voting power a priori. The first is the Bernoulli model which is at the heart of the Penrose measure; the other is a rival that leads to the Shapley-Shubik index. Even though the second measure is often introduced axiomatically,[6] it can be regarded as the probability of pivotality under a model different from the Bernoulli model (Theorem 6.3.13 on p. 208 in Felsenthal and Machover 1998).[7] A clarification of the probabilities in

[5] I will here assume that probabilities are ascribed to events from an event space. An alternative option is to assign probabilities to propositions, see e.g. (Howson and Urbach 2006, pp. 13–14).

[6] See (Felsenthal and Machover 1998, Chap. 6).

[7] Felsenthal and Machover (1998, p. 210) deny that the Shapley-Shubik index is a priori because it is not based upon the Principle of Indifference. However, the principle has come under attack because it does not lead to unambiguous results in some cases (see Gillies 2000, pp. 37–49 for a textbook account of the principle and its problems). Further, it is arguable that the principle leads to the Shapley-Shubik index provided that the space of ultimate possibilities is re-defined (see below; see Mellor 2005, pp. 24–26 for ultimate possibilities).

voting theory may be helpful to find out how the values of a priori measures of voting power should be set.

There are other reasons why the interpretation of probability is important for our understanding of voting power. As has already been mentioned, it is often suggested that voting systems are preferable if they allot equal voting powers to the citizens. But why are they so? What is the normative significance of voting power? This question can only be answered if we have a clearer understanding of the related probabilities.

I do not know of any systematic discussion of the question of this paper in the existent research literature. Even Morriss (1987), who provides a detailed philosophical analysis of power and who proposes to use probabilities to measure voting power, does not say much about what the probabilities are. But my paper can draw on an extensive philosophical literature about probabilities in general. Of course, not every argument from this literature will be relevant because our concern is not probability in general, but just one special probability.

The method of this paper is to work through a list of well-known interpretations of probabilities. For each interpretation, I will check whether it is adequate for understanding voting power and I will trace the consequences for the way the values of the probabilities should be set. I begin with a very natural suggestion, viz. that the probability of pivotality measures the strength of a disposition (Sect. 2). I nevertheless reject this interpretation as an understanding of voting power, and I move to other objectivist interpretations of probabilities in Sect. 3. Subjectivist readings of probabilities can quickly be dismissed (Sect. 4). While the argument in Sects. 2–4 is independent of the uses to which measures of voting power are put, I turn to such uses in Sect. 5. It turns out that most objective probabilities do not underwrite the measurement of voting power for normative purposes in the way people seem to hope. As a solution, I suggest in Sect. 6 that we understand the probabilities in voting theory as classical probabilities. I draw my conclusions in Sect. 7.

This paper is restricted to binary voting games, i.e., every voter has only two options and abstention is not possible. Monotonicity will also be taken for granted.[8]

2 Voting Power as the Strength of a Disposition

(Felsenthal and Machover 1998, p. 36) claim that probability is the obvious candidate for quantifying the degree to which a voter can make a difference in a collective vote. Some people may disagree because degrees are not often measured using probabilities. For instance, the degree to which an object looks blue is not the probability that the object looks blue; and the degree to which a particle is positively charged is not measured using the probability that the particle is positively charged.

[8] See Felsenthal and Machover (1998), Definition 2.1.1 on p. 21 for the definition of monotonicity.

But even though the step from a determinable property to a probability is less than obvious in general, things may be different when we turn to dispositional properties. The idea is that dispositions can come in various strengths and that probabilities measure these strengths. This idea is often used to explain what an objective probability is. Now voting power seems clearly dispositional.[9] This leads to the following suggestion:

> As a measure of voting power, the probability of pivotality is the strength of a dispositional property.

This suggestion is meant to suffice for spelling out what the probability of pivotality is.

The suggestion is natural, and it would explain why Felsenthal and Machover (1998) think that probability is arguably the only way to measure the degree to which a voter can be decisive. Taking for granted the suggestion, the step from power to probability draws only on a definition of probability. Under several alternative interpretations of probabilities, the step from power to probability is less trivial.

But there seem to be problems about the suggestion. Some of them can be solved, while others are more difficult to deal with.

Dispositions and dispositional properties always attach to *an object that bears the property*, and they are dispositions *to do something*, call it φ, *in certain circumstances*. φ encodes the effect that becomes manifest; the circumstances are called manifestation conditions. For instance, the fragility of a glass is a dispositional property *of this very glass*, and the relevant disposition is the tendency of the glass *to break* if *dropped in a certain way*. The strength of this disposition may then define the probability that the glass breaks if dropped.[10]

If the probability of pivotality is the strength of a disposition, it may seem that the object to which the disposition attaches is a voter because the voter is said to be pivotal with a certain probability.[11] But it does not make sense to say that a voter has the dispositional property of voting power because the extent to which a voter can make a difference does not so much depend on this very voter and her properties, but rather on the other voters and the voting system with its voting rule. In philosophical parlance, this is to say that the categorical basis of the disposition includes properties of the other voters and of the voting system. Thus, as a dispositional property, the probability that a voter is pivotal cannot be ascribed to the voter, but should rather be ascribed to a setting that includes the other voters and the voting rule.

The point is familiar from other examples. For instance, coins are often said to have a probability of landing heads. This suggests that the probability attaches to the coin. But the categorical basis of the related disposition is broader than the coin

[9]For example (Morriss 1987, p. 19).

[10]See Fara (2009) for an introduction to dispositions.

[11]Voting powers are sometimes ascribed to votes, and at other times to voters. This will not make a difference in what follows, and, for simplicity, I will always assign voting power to voters.

itself; it includes e.g. properties of the person that flips the coin. Thus, the disposition is more appropriately said to attach to a chance setup.[12]

Another question is what the manifestation of the disposition is like. The answer is not trivial because voting power is defined as the extent to which a voter can make a difference to the outcome of a collective decision. This specification has still a modal notion in it ("*can* make a difference") and is thus markedly different from the specification of other dispositions. For instance, fragility is the disposition to break, and not the disposition that something *can* break.

The problem can easily be solved though. In the setting usually assumed in voting theory, a voter can make a difference if, and only if (iff), she is pivotal. This is so iff the outcome of the election would be different had the voter voted differently. This counterfactual is always understood in the following way. If S is the coalition of yes-voters, voter a is pivotal iff $S \setminus \{a\}$ is losing, while $S \cup \{a\}$ is winning.[13] But in this situation, the voter *can* not only make a difference, but *does* in fact make a difference. As a consequence, what the manifestation of the disposition under consideration is can be described without recourse to modal notions. My conclusion is that there is nothing problematic about the disposition itself.[14]

An analysis of probability in terms of a strength of a disposition would nowadays count as a propensity view of probability, at least if we adopt a terminological suggestion made by Gillies (2000, p. 126). Gillies (2000, p. 126) distinguishes two types of propensity theories, which he calls long-run and single-case propensity theories.[15] In these terms, we are now talking about a single-case propensity theory simply because the disposition does not essentially refer to a series of events. Under long-run propensity theories, by contrast, the crucial disposition is a disposition to produce relative frequencies of a certain value. The crucial distinction turns on the question of what exactly the value of a probability is. Under a single-case propensity view, it reflects the strength to which a single event may happen; under the long-run propensity view, it is the value of a relative frequency that will become manifest if the manifestation conditions are fulfilled in a series of trials.[16]

[12]There is a philosophical debate about what exactly probabilities qua strengths of dispositions attach to. Some have suggested that, properly speaking, the related chance set-up includes the whole world. See (Gillies 2000, pp. 126–129) for an overview of corresponding positions.

[13]See Felsenthal and Machover (1998), Definitions 2.3.4 and 2.3.6 on pp. 24–25.

[14]This is not to reject the general distinction between power and influence, which is rightly stressed by Morriss (1987, Chap. 2). I only think that the distinction crumbles if we turn to voting power. There is an important difference to other sorts of powers at this point. I can have the power to play Beethoven's Pathétique on the piano, but simply decide not to execute the power. But if I have the power to make a difference in a collective decision, I cannot decide not to execute this power.

[15]See also (Eagle 2004, pp. 377–383) for a more refined classification of propensity theories.

[16]In the terms of Eagle (2004), my discussion of a single-case propensity theory is restricted to what Eagle calls tendency accounts (p. 379). I thus bracket the distribution display account attributed to Mellor. This account is close to Lewis's account, which will be discussed below.

In the philosophical discussion, single-case propensity views have not fared well and they are not often defended. One main reason is that the propensities with their strengths are merely postulated. A single-case propensity theory thus has metaphysical "costs". But its benefits are modest. It does not explain why probabilities obey the probability calculus and why they are empirically determined in the way they are. For if probabilities are strengths of dispositions to produce certain events in a single case, they are quite elusive. It is not clear why they should obey the rules of the probability calculus,[17] and how we should be able to fix the values of such probabilities. In practice, statistical data and thus relative frequencies are supposed to provide evidence for probabilistic claims. Proponents of a single-case propensity view would have to postulate certain assumptions about their propensities to make sense of this. See Eagle (2004) for a number of similar objections to single-case propensity views.

I do not want to conclude that single-case propensity theories of probability are hopeless, but I think that there are good reasons to avoid them if our task is to understand voting power. Single-case propensity theories have a number of general problems, and why let them affect our understanding of voting power? In particular, single-case propensity raise a number of metaphysical questions, but why should metaphysics be so important for voting theory?

This is not to deny that there are powers or dispositional properties. Nor is this meant to negate that voting powers are in some sense properties of voting systems (or of larger set-ups). What is worrisome is only the idea that dispositions come in degrees and that these degrees fix the meaning of the probabilities.

It is interesting to compare to the approach by Morriss (1987) at this point. Morriss takes power to be a dispositional property (e.g. p. 19). He also uses probabilities to measure powers (Chaps. 22–23). But this does not commit him to say that the pertinent probabilities are degrees of dispositions. At least, I am not aware of a hint that he defines probabilities as degrees of dispositions.

3 Other Objectivist Interpretations

My suggestion then is to resist the temptation to explain our probabilities as strengths of dispositions. Nevertheless, the probabilities may be understood in an objectivist way. Under objectivist accounts, probabilistic statements have truth conditions that do not refer to human attitudes and needs. If some such statements are true, there are objective facts as to what values the probabilities under consideration take. Probabilities that are interpreted as objective are often called chances. In this section, I will briefly examine a few other objectivist views of probabilities and ask

[17] See (Eagle 2004, pp. 384–385) for this criticism.

whether they provide a useful reading of the probabilities that are used to quantify voting power.[18]

The simplest objectivist view of probabilities is *actual frequentism*.[19] Under actual frequentism, probabilistic statements collapse with statements about relative frequencies, or probabilities are relative frequencies in actual series of events. Actual frequentism is a plausible interpretation for the statement that each Roman citizen has a probability of 0.3 to own a cat. This is presumably no more than saying that 30 % of the Roman citizens own a cat.

Actual frequentism is not plausible if we are to understand measures of voting power. One reason is that, under actual frequentism, most voters would have zero voting power for most voting systems because most actual voting systems never produce a situation in which voters are pivotal. But voting theorists do not want to say that most people have mostly zero voting power. If the citizens had almost always zero voting power, then voting power could not discriminate between different voting systems. Further, power clearly is about what can or may happen, and not just about what happens.

Problems with actual frequentism can be avoided if we identify probabilities not with actual frequencies, but rather with *hypothetical frequencies* that would arise if a certain experiment of chance were repeated. This idea is at the center of *long-run propensity views*. However, there are a number of difficulties to fill in the details of this view quite generally. For instance, how often are we to repeat a chance experiment (the collective vote in our case) to obtain the probability? The value of a probability will certainly depend on the number of trials unless we require there to be infinitely many trials. But an infinite series of trials has its own problems; for instance, the value of the probability may depend on the order in which the trials are evaluated.[20] Another question is whether there are any facts as to what relative frequencies of trials would be in an infinite series of trials.[21] These difficulties are also relevant to the measurement of voting power.[22]

A more promising objectivist account of probabilities is the *Humean account* by the later D. Lewis.[23] The account allows one to assign probabilities to single events, but the values of these chances are fixed using regularities in the pattern of actual events. Very briefly, the chance of an event is the probability P' that a best system of the whole world would ascribe to the event. The best system is the winner in a competition among systems of sentences about the world. The sentences are

---

[18]Some authors call probabilities objective iff a weaker condition is fulfilled, viz. that their values are uniquely fixed for rational persons. See e.g. (Uffink 2011, pp. 25–26) for this point. In this paper, I use the stronger notion of objectivity.

[19]See Hájek (1997) for a discussion of this position.

[20](Hájek 2009, pp. 218–220).

[21](Hájek 2009, pp. 217–218).

[22]See Hájek (2009) and Eagle (2004) for criticism of views that appeal to hypothetical frequencies.

[23]See Lewis (1994) and consult Lewis (1980) for an important fore-runner. For recent appraisals see Loewer (2004), Hoefer (2007), Frigg and Hoefer (2009).

allowed to have an uninterpreted probability function, which I have called P'. The best system achieves an optimal balance between strength (informativeness about the world), simplicity and fit. A system fits the world the better, the higher the P' of the real world is. Axioms or theorems of the best system that refer to P' (or to chances) are called chance laws.

This account can be illustrated as follows. Scientists have a hard time to provide a deterministic theory of the outcomes of experiments with electron spins. They can nevertheless define a P'-function for the outcomes of certain types of experiments. This function should be very simple, but also return comparatively large values for the patterns in the real world. The best P'-function that scientists can come with is then taken to be the chance function, and it implies the values of probabilities of single trials.

We can try to apply this idea to voting. The hope is that actual votes in the world display patterns that are optimally captured using one or several simple probability models. There may be one probability model over all votes or several models that cover votes under different circumstances. In either case, the probability that a voter is pivotal may be different from the actual relative frequency with which she is pivotal, simply because we may gain a lot of simplicity when we allow for probabilities that deviate from the actual relative frequencies.[24]

It may turn out though that the best system of the world does not assign probabilities to votes and voters being pivotal. To check whether this is so or not inquires empirical investigation and goes beyond the scope of this paper. But this not a shortcoming specific to the Humean approach. The point applies more generally to all kinds of objectivist accounts. Whether there are objective probabilities over a certain range of events or not is subject to empirical scrutiny and cannot be decided from the philosophical arm chair. What we can argue though is this: It is conceivable that the numbers used in voting theories are objective probabilities in a specific sense. If certain facts turn out to be right, as it were, the probabilities used in voting theory are such objective probabilities. The arguments that I have leveled against the propensity view earlier in this paper cast doubts on the very idea that the numbers from voting theory may be objective probabilities in the sense of this very view.

Are there similar doubts concerning a Humean view of the probabilities? As N. Hall and N. Bostrom have noted, Lewis's criterion of fit does not work as required when applied to infinite spaces of events (Elga 2004). However, the event space of a single vote is finite, so the problem arises only if there is an infinite sequence of votes in the actual world. And even for this case, a solution to the problem has been proposed by Elga (2004). I cannot discuss this proposal in detail and conclude that there are at least some reasons to think that, for our purposes, the zero-fit problem pointed out by Hall and Bostrom doesn't spoil everything.

Lewis himself voices doubts as to whether his account is really objectivist (Lewis 1994, p. 479). The problem is that it draws on the notions of simplicity and informativeness, and judgments to the effect that a system is simple or informative

[24] Cf. (Lewis 1994, p. 481).

may not reflect mind-independent facts. However, for the purposes of this paper, we should not worry too much about this issue. For one thing, full mind-independence is not a prospect in voting theory because we are interested in human behavior. For another thing, we are not here doing metaphysics, and stable probabilities the values of which people can rationally agree upon on the basis of data may be sufficient for voting theory.

To conclude this section: There is at least one objectivist account, viz. the Humean account that may explain what our probability of pivotality is. For the purposes of this paper, an even weaker claim is sufficient: There are objectivist candidate interpretations that may account for the probability of pivotality in voting theory. But what about subjectivist views? In the next section I turn to such views.

4 Subjectivist Interpretations

Subjectivist accounts of probability have a number of merits and seem a promising start to understand many probabilistic statements. In particular, they do not raise metaphysical questions. In the following, I will concentrate on one very elaborate subjectivist account. My argument can be generalized to other subjectivist accounts as well.

Under the account to be considered, probabilities are degrees of belief. There are reasons to think that belief comes in degrees, and there are proposals how to measure actual degrees of belief. The broad idea is that beliefs and their strengths manifest themselves in the (hypothetical) behavior of a person. They influence how much money a person would bet on a certain event, for instance.[25]

However, in voting theory, subjectivist accounts are non-starters. When voting theorists calculate measures of voting power, they are interested in objective features of voting systems. They are interested in the degree to which a voter can make a difference and not in the degree to which somebody believes or may rationally believe that a voter makes a difference.

To put the same point in different words: Each subjective view of probabilities in voting theory renders measures of voting power pure estimates. These estimates may either be estimates of objective probabilities or refer to non-probabilistic matters of fact. In the first case, ultimately only the objective probabilities seem to be of real interest and fundamental for the understanding of voting power. In the latter case, our measures of voting power are degrees of belief that we assign to matters of non-probabilistic fact, e.g. to the event that a voter is pivotal in a particular decision. This event will either occur or not, and the voter will either be pivotal or not. If we do not

[25] See Ramsey (1931) and de Finetti (1931a), de Finetti (1931b), de Finetti (1937) for important original contributions, and (Gillies 2000, Chap. 5) and (Mellor 2005, Chap. 5) for textbook accounts.

know whether she will be pivotal and put a probability on this event, this is our own business. We cannot say that it quantifies a feature of the voting system.

This is not to deny that many claims about probabilities of pivotality reflect guesses and estimates. Under most views of objective probabilities, the latter are not observable and not easily determinable otherwise, and our best efforts will only lead to estimates that will not normally coincide with the true values of the probabilities. But even if our probabilistic statements do not reflect the truth of the matter, our intention is still to refer to objective probabilities.

Our conclusions so far can be summarized as follows. The degree to which a voter has power to affect the outcome of a collective vote cannot be a subjective probability. It must thus be objective. There is at least one objectivist view of probability that may account for the probability of pivotality under consideration.

5 Uses of Voting Power

So far, my argument has only been constrained by the general idea that voting power is the extent to which a voter can make a difference. I have abstracted from the uses to which measures of voting power are put. But the purposes for which measures of voting power are used constrain the choice of an interpretation of the probabilities, or so I shall argue.

We can distinguish between *descriptive* and *normative* uses of measures of voting power.[26] We use measures of voting power descriptively when our aim is to state how much power a voter has in a certain voting system and thus to convey information about the voter and the system. We use measures of voting power normatively if our aim is to normatively assess alternative voting systems. If we do so, we assume that voting power has normative significance and that it figures centrally in a normative principle. This principle may require that the voting powers of all citizens be equalized.

Descriptive uses of measures of voting power can be dealt with very quickly. Some objectivist interpretations of probabilities will do for such uses, provided that there are no general problems for the interpretations. The reason is that, under an objectivist interpretation, probabilistic statements concern matters of fact, and every statement about matters of fact can be used to convey information and to characterize an object (a voting system and voters in our case).

Things are different when we turn to normative uses of voting power. As a matter of fact, when voting theorists calculate measures of voting power to normatively assess a voting system, they most often measure voting power a priori. That is, they do not take into account empirical information, but rather adopt the Bernoulli model. In fact, attempts to equalize realistic probabilities of pivotality seem odd.[27]

---

[26] Cf. (Morriss 1987, Chap. 6).

[27] See (Felsenthal and Machover 2000, p. 13) for the expression of related worries.

Suppose, for instance, that five voters, a–e, take collective decisions following simple majority voting, i.e., a proposal is accepted iff there are at least three votes in favor of it. Assume further that, with a high probability, three voters, a, b and c, vote exactly the other way d and e do. As a consequence a, b and c are pivotal with a probability close to one, while d and e are pivotal with a probability close to zero. But does this mean that we should change the voting system? Should we give d and e additional votes to enhance their probability of pivotality? This seems odd.

But if measures of voting powers are objective probabilities of the kind discussed so far, then it is a matter of facts what their values are, and empirical data should be used to determine the values of the measures. The reason is that the most promising objectivist interpretations discussed so far claim a conceptual link between relative frequencies and probabilities. According to the long-run propensity view, probabilities measure propensities to produce certain frequencies, and we should certainly use actual frequencies to learn about the propensities. According to the Humean conception by Lewis, chances have to fit the patterns of actual events in the world, which implies once more that data should be used to constrain the values of the probabilities.

This leads to the following problem: On the one hand, voting theorists abstract from empirical information if they calculate powers to assess voting systems. On the other hand, as probabilities, measures of voting power seem well understood as objective chances, but such an understanding of the probabilities pushes us towards the use of empirical information, which has counterintuitive implications.

As a reaction, voting theorists have tried to find arguments explaining why normative assessments of voting rules should be based upon the Bernoulli model and not take into account empirical data. The idea could be as follows. Measures of voting power are objective probabilities, but when it comes to normative assessments of voting rules, additional considerations push us towards an a priori model. In my view, many arguments proposed thus far fail.

According to the *first* argument, preferences are too malleable to be foreseen in the long run. Thus, votes, which are based upon preferences, cannot be predicted. Consequently, past data about votes cannot be used to estimate the probabilities over voting profiles in the future. But normative assessments of voting rules typically concern the future. As a consequence, we should not use data when calculating measures of voting power to normatively assess a voting rule. It is more reasonable to use the Principle of Indifference to fix the values of the probabilities in an a priori manner. This argument can be found in (Felsenthal and Machover 2000, p. 13). Call it the epistemic argument.

The epistemic argument is not very convincing. It is certainly not possible to predict preferences and votes with certainty, but empirical data are most often a better guide to the probabilities over voting profiles than guesses based upon an a priori model.[28] Felsenthal and Machover (1998, p. 21) admit this when they claim

[28] See Gelman et al. (2004) for empirical work about the probability of pivotality.

that an assessment of a voting system may draw on a posteriori information about preferences if "long-term systematic real factors" are known.

According to a *second* argument, we are here really concerned with a *moral* assessment, and the latter is subject to a well-known constraint, viz. a Rawlsian "veil of ignorance". The veil prohibits the use of certain information, and it then claimed that the veil excludes knowledge about preferences (Felsenthal and Machover 2000, p. 13). As a consequence, empirical data about votes should not be taken into account.

I agree that the veil of ignorance is an important device in moral philosophy. Moral judgments are impartial, and an impartial perspective prohibits the use of some information, particularly of information about the position one will take up in society and about *one's own* preferences. In Rawls' words,

> [. . .] it should be impossible to tailor principles to the circumstances of one's own case. We should insure further that particular inclinations and aspirations, and persons' conceptions of their good do not affect the principles adopted (Rawls 1971, p. 16).

But does it follow that one is not allowed to "tailor the decision rule to the specific interests, preferences and affinities of the voters" (Felsenthal and Machover 2000, p. 13, my emphasis)? This is not at all obvious. Rawls himself explicitly does not permit people to use information about the "particular circumstances of their society" (p. 118). But later he proposes a more detailed prescription for creating a just political system. It is called a "four-stage sequence". Very roughly, one starts with finding general principles and moves on to apply these principles to more and more concrete settings. As one moves to more concrete problems, the veil of ignorance is lifted stepwise (p. 172). For my purposes, I can focus on the first two stages. The point of the first stage is to identify the basic principles of justice. At the second stage the task is to choose a "constitutional convention". This convention is supposed to guarantee "equal citizenship" (p. 173). It is arguable that the specification of a voting scheme is part of this task, and that the desideratum of equal voting power spells out the idea of equal citizenship. But what kind of information is admitted at this stage? According to Rawls, people are allowed to take into account "the relevant general facts about their society, that is, its natural circumstances and resources, its level of economic advance and political culture, and so on" (p. 172). One may argue that information about the distribution of preferences of votes is "a general fact about society" and could be placed under the "and so on"-clause. It is also clear that knowledge about the general pattern of preferences is not the kind of information that individuals could use to tune the choice of a constitution to their own advantage. Hence, one might argue, general information about preferences is not excluded by the veil of ignorance.

Third, Laruelle and Valenciano (2005, p. 183) argue that one should abstract from the voters' preferences if a normative assessment is focused on a voting scheme as such. This sounds quite right. But it is questionable whether this argument applies to most normative assessments of voting rules. For instance, Felsenthal and Machover (2000) consider the Council of the European Union and discuss various possible voting rules. In this application, they are not concerned with a voting rule as such, but rather with a voting rule for the European Union.[29] It is furthermore possible that a voting system as such need not have well-defined probabilities over voting profiles. Recall our point above that, properly speaking, some probabilities attach not to objects but rather to chance setups.

A *fourth* and yet different argument to the effect that, in the assessment of a voting system, the aprioristic Bernoulli model should be chosen can be extracted from Morriss (1987). When Morriss considers the design of two-tier voting systems in his Chap. 22, he restricts himself to ability rather than to ableness (pp. 183–184, see Chap. 11 for the distinction between ability and ableness). Assignments of ability concern what a person can do quite generally and independently of her opportunities. The idea is further, very roughly, that you have more ability to do something than me if you can do that under more possible circumstances than me. In Morriss's view, the possibilities are to be weighted according to their importance, and he does not see any reasons to say that one voting profile matters more than does another (p. 159). Thus, each possible voting profile has the same probabilistic weight, and we end up with the Bernoulli model.

I am sympathetic to Morriss's views, but I think that the argument just mentioned has loopholes. First, why should a normative assessment of voting systems focus on ability rather than ableness? Second, when we quantify ability, why can't we say that some possibilities are more important than others, simply because they are realized with a higher probability? Further, it is clear that the argument in the previous paragraph moves beyond the range of objective probabilities discussed so far.

The upshot is that we have not yet found a convincing argument to the effect that we should not use empirical information to measure voting power for normative purposes. But the objectivist interpretations of probabilities considered so far push us to use empirical information to set their values. To underwrite the use of a priori probabilities in normative assessments of voting theory, we need a different interpretation of the probabilities.

6 Powers and Rights

To address this task, it is useful to step back a bit. The most prominent normative use of measures of voting powers is based upon the idea that citizens should have equal voting powers. But why should they? What is the point of equalizing voting power?

Arneson (2007, p. 593) draws a useful distinction between "equality of democratic citizenship" and "equality of condition". The former concerns freedoms and

[29]When Laruelle and Valenciano (2005) explain what they mean by the "*normative* point of view" (183), they say that a related assessment is concerned with a voting situation or rule, "irrespective of what voters occupy the seats." It may be objected that this is too narrow a conception of "normative".

rights, while the latter is about the actual conditions in which people live (Arneson 2007, p. 594). According to Arneson, the latter is “an amorphous ideal”, which “cries out for clarification” (Arneson 2007).

If the equalization of voting power is to be an ideal, it plausibly requires equality of democratic citizenship in the terms of Arneson. To vote is to use a basic democratic right, and equal rights are a matter of equal citizenship.

But what exactly does equality of democratic citizenship require of voting rules? In his “Theory of Justice”, Rawls writes

> The principle of equal liberty, when applied to the political procedure defined by the constitution, I shall refer to as the principle of (equal) participation. It requires that all citizens are to have an equal right to take part in, and to determine the outcome of, the institutional process that establishes the laws with which they are to comply. (Rawls 1971, p. 194).

In this passage, Rawls does not just require that citizens have the right to participate at votes, but also that their rights to determine the outcome are equal. This sounds plausible, but what does it mean to say that each citizen deserves an equal right to determine the result of the legislative process? I take it that a voter determines the outcome of a vote iff she is pivotal. But Rawls cannot mean to say that every citizen must always be pivotal because pivotality is rare if people freely use their right to vote the way they want. The most straight-forward way to make sense of Rawls’s postulate is thus to require that citizens have the same *possibilities* to make a difference to the outcome. This is to say that for each citizen there should be the same number of possibilities (i.e., possible voting profiles) in which she is pivotal.

Why do I say that citizens should have the same *possibilities* to determine the outcome? I say so because rights are about possibilities. A person has a right to do ϕ iff she can do ϕ without having to expect interference. When we fix rights, we think about possibilities. Rawls’ liberty principle (Rawls 1971, p. 53, 220) requires that equal liberties or rights are compatible with each other. Two rights are incompatible if the rights cannot be used at the same time. Thus, when we fix rights, we think about possibilities and not about what people are most likely to do.

There is a long tradition according to which probabilities measure possibilities. According to the *classical theory/interpretation*, the probability of an event is the number of ultimate possibilities compatible with the event normalized by the total number of such possibilities.[30] Thus, under the classical theory, the requirement that each citizen have the same possibility to determine the outcome of a collective decision boils down to the postulate that each citizen have the same probability to do so. Thus, equalizing a classical possibility is exactly what voting theorists do when they equalize the Penrose measures of all voters because this measure assumes that every possible voting profile has the same chance of occurring.

This suggests that we can justify the demand that voting power be equalized under an a priori measure of voting power if we use classical probabilities

[30]Consult Gillies (2000, Chap. 2) and Mellor (2005, Chap. 2) for this interpretation.

to measure voting power. The classical interpretation thus makes sense of the normative uses that are made of measures of voting power. Note that the classical interpretation is in some sense objectivist, so we are not caught in the same difficulty that I have raised for subjectivist interpretations. The classical interpretation has another virtue. It explains the validity of the axioms of the probability calculus. The axioms have not to be postulated on top of the classical view, but rather follow from it.[31]

Nevertheless, the classical interpretation is not very popular these days and it is sometimes regarded as outdated. As it happens, at least to some part, general criticism of the classical interpretation does not apply in the field of voting theory.

One problem is that the classical view implies a priori that probabilities of rival events are equal at some level.[32] This is certainly a problem if we are interested in probabilities as they are used in science because we would like to allow for the possibility that probabilities of rival events are not equal even at a very fundamental level.[33] But we need not worry about this problem in our case because, in normative uses of voting power, there is no motivation to allow for unequal probabilities at the most fundamental level.

Another problem about the classical view is that it has the Principle of Indifference built into it. This principle is often thought to be problematic because it can lead to paradoxes.[34] Depending on how exactly the sample space is described, different probability distributions arise from the principle. To be sure, only the probabilities of *ultimate* possibilities should be equal,[35] but the question is what the ultimate possibilities are. In science, this problem may be decided using empirical data, but this will not do for our purposes.

Is this problem relevant in voting theory? We have a finite sample space (i.e., the different voting profiles), and, at least at first sight, there seems only one natural option to assume equal probabilities on it, viz. to give each voting profile the same probability, as it is done under the Bernoulli model. In particular, the classical theory does not seem to underwrite the choice of a rival of the Penrose measure, viz. the Shapley-Shubik index. This index assumes that each possible number of yes-votes has the same probability. Different ways to realize a given number of yes-votes are then again given equal probabilistic weights.[36] Although the equalization of probabilities is crucial for this model, the probabilities do not arise by equalizing the probabilities of ultimate possibilities.

It is true though that we obtain the Shapley-Shubik index as a measure of voting power if we re-define the sample space and the ultimate possibilities. The idea would be to say that each voting profile does not have the same probability because voting

[31] This is a contrast with some propensity views, see (Mellor 2005, p. 25) for some details.

[32] See (Mellor 2005, pp. 25–26) for a discussion.

[33] See (Mellor 2005, pp. 25–26) though.

[34] See (Gillies 2000, pp. 37–49) again.

[35] (Mellor 2005, p. 24).

[36] See Felsenthal and Machover (1998), Remark 6.3.12(ii) on p. 207.

profiles are not ultimate possibilities. Rather, most profiles would contain several ultimate possibilities. The latter would have to take into account orders in which the yes- and the no-votes arise, respectively.[37] Thus, under this count, there would be two possibilities under which a votes yes and b votes yes.

If there are in fact two sensible ways to count possibilities, then the idea of a unique classical probability crumbles. In any case, there would be a problem for normative assessments of voting rules, because equalizing different counts could favor different voting systems. The impact on the debate of how a priori power should be measured (second issue in the introduction) is immediate.

My hope is that we can rule out one of the proposed ways to count possibilities. In fact, the sample space and the identification of ultimate possibilities under the Shapley-Shubik index seem unnatural and unwarranted. It is even questionable whether we can make sense of the idea that the votes arise in different orders. And even if we can, why should order matter for ultimate possibilities? Maybe, considerations about order can be ruled out, when we take into account practical concerns about rights. If this is so, there is only one sensible way to identify ultimate possibilities, we have uncontroversial classical probabilities, and we should use the Bernoulli model and thus the Penrose measure.[38]

I conclude that classical probability is a promising route to understand probabilities in voting theory. The reason is that the concern about equal voting powers is about rights or legal powers and that arguments about rights appeal to what is possible and not to what is likely as a matter of fact. This suggests that measures of voting powers should count possibilities. Under the classical interpretation, such counts become probabilities, and we can thus say that the probabilities in the Penrose measure are classical probabilities. But admittedly, my argument about rights and possibilities is still in need of precisification and I should be able to say more about why real-world chances don't matter for rights. Further, I'd like to have a stronger case that there are in fact unique classical probabilities over votes.

7 Conclusions

When we quantify voting power using a probability, what type of probability are we dealing with? Which interpretation of probability is most fitting? I have argued that the answer to this question depends on the use to which a measure of voting power is put. If the aim is to describe the voting powers of voters most realistically, then any latter-day objectivist interpretation (for instance that of Lewis) will do, provided there are any probabilities of the sort suggested. Things are different when we

[37]Consult Theorem 6.3.13 on p. 208 in Felsenthal and Machover (1998) for the mathematical basis of this argument.

[38]This is not to deny that the Shapley-Shubik index may succeed as a measure of P-power. See (Felsenthal and Machover 1998, Chap. 6).

turn to normative uses of measures of voting power. Currently discussed objectivist interpretations would push us to use empirical information to set the values of the probabilities, but this does not seem fitting when we normatively assess a voting system. Nor do subjectivist interpretations fit the bill. I have thus suggested that the classical view of probabilities is most appropriate to make sense of the probabilities in voting theory. Under this interpretation, probabilities count possibilities. Such a count of possibilities seems appropriate when we equalize the voting powers of different voters. For to equalize voting power is a requirement of equality of democratic citizenship; and this type of equality is about rights. Now when we think about rights, we are thinking about possibilities. For instance, rights are supposed to be compatible with each other. This means that several parties can possibly do what they have a right to do according to the law. A potential problem with the classical interpretation is that there may be different ways to count possibilities. In voting theory, one way of counting possibilities, viz. that underlying the Bernoulli model, seems most appropriate, but more research is needed to argue this point more forcefully.

The results of this paper have two interesting consequences. First, they suggest for the philosophy of probability that the classical view, which is not that fashionable these days, can have its merits in a suitable context. Second, there is a consequence for voting theory. To switch between a priori and a posteriori measures of voting power is not just a move from one probabilistic model to another one. Rather, the meaning of the probability changes. While a posteriori measures are plausibly taken to quantify power as a matter of condition in the terms of Arneson, a priori measures are more fittingly taken to be about certain possibilities, which are important in the realm of rights.

Acknowledgements This is a wonderful opportunity to thank Dan Felsenthal and Moshé Machover. I owe them a lot, and their advice and support has always been helpful to me. Thanks also to the editors of this volume and to an anonymous referee. Some parts of this paper draw on joint work with Luc Bovens to whom I am grateful.

References

Arneson, R. (2007). Equality. In R. Goodin, P. Pettit, & T. Pogge (Eds.), *A companion to contemporary political philosophy* (pp. 593–611). Oxford: Blackwell.

Beisbart, C. (2010). Groups can make a difference: Voting power measures extended. *Theory and Decision, 69,* 469–488.

Beisbart, C., & Bovens, L. (2008). A power measure analysis of amendment 36 in Colorado. *Public Choice, 124,* 231–246.

Bovens, L., & Beisbart, C. (2011). Measuring influence for dependent voters: A generalisation of the Banzhaf measure. *Synthese, 179,* 35–56.

de Finetti, B. (1931a). Probabilismo [Probabilism]. *Logos, 14,* 163–219 [Translated as de Finetti, B. (1989). A critical essay on the theory of probability and on the value of science. *Erkenntnis, 31,* 169–223].

de Finetti, B. (1931b). Sul significato soggetivo della probabilità. *Fundamenta Mathematica, 17,* 298–329.

de Finetti, B. (1937). La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré, 7,* 1–68 [English translation as: de Finetti, B., Foresight: Its logical laws, its subjective sources. In H. E. Kyburg & H. E. Smokler (Eds.), *Studies in Subjective Probability*, pp. 53–118, 1964, New York: Wiley].

Eagle, A. (2004). Twenty-one arguments against propensity analyses of probability. *Erkenntnis, 60,* 371–416.

Eagle, A. (Ed.). (2011). *Philosophy of probability. Contemporary readings*. London, England: Routledge.

Elga, A. (2004) Infinitesimal chances and the laws of nature. *Australasian Journal of Philosophy, 82*, 67–76.

Fara, M. (2009). Dispositions. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (summer 2009 ed.).

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

Felsenthal, D. S., & Machover, M. (2000). *Enlargement of the EU and weighted voting in its council of ministers*. Retrieved from http://eprints.lse.ac.uk/407/,2000

Frigg, R., & Hoefer, C. (2009). Determinism and chance from a Humean perspective. In D. Dieks, W. Gonzalez, S. Hartmann, M. Weber, F. Stadler, & T. Uebel (Eds.), *The present situation in the philosophy of science*. Berlin/New York: Springer.

Gelman, A., Katz, J. N., & Bafumi, J. (2004). Standard voting power indexes don't work: An empirical analysis. *British Journal of Political Science, 34,* 657–674.

Gillies, D. (2000). *Philosophical theories of probability*. London/New York: Routledge.

Hájek, A. (1997). 'Mises Redux'– Redux. Fifteen arguments against finite frequentism. *Erkenntnis, 45,* 209–227.

Hájek, A. (2009). Fifteen arguments against hypothetical frequentism. *Erkenntnis, 20,* 211–235.

Hájek, A. (2010). Interpretations of probability. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (spring 2010 ed.). http://plato.stanford.edu/archives/spr2010/entries/probability-interpret/

Hoefer, C. (2007). The third way on objective probability: A sceptic's guide to objective chance. *Mind, 116,* 549–596.

Howson, C., & Urbach, P. (2006). *Scientific reasoning: The Bayesian approach* (3rd ed.). La Salle, IL: Open Court.

Kaniovski, S., & Leech, D. (2009). A behavioural power index. *Public Choice, 141,* 17–29.

Kolmogorov, A. N. (1956). *Foundations of the theory of probability* (2nd ed.). New York: Chelsea.

Laruelle, A., & Valenciano, F. (2005). Assessing success and decisiveness in voting situations. *Social Choice and Welfare, 24,* 171–197.

Lewis, D. (1980). *A subjectivist's guide to objective chance*. In R. C. Jeffrey (Ed.), *Studies in inductive logic and probability* (Vol. 2). Berkeley, CA: University of California Press [Here quoted from the reprint in Lewis (1986), pp. 84–113].

Lewis, D. (1986). *Philosophical papers* (Vol. 2). New York: Oxford University Press.

Lewis, D. (1994). Humean supervenience debugged. *Mind, 104*, 473–490 [Reprinted in D. Lewis, *Papers in metaphysics and epistemology*, 1999, Cambridge: Cambridge University Press].

Loewer, B. (2004). David Lewis's humean theory of objective chance. *Philosophy of Science, 71*, 1115–1125.

Machover, M. (2007). *Discussion topic: Voting power when voters' independence is not assumed*, Mimeo. http://eprints.lse.ac.uk/2966/

Mellor, H. D. (2005). *Probability. A philosophical introduction*. London and New York: Routledge.

Morriss, P. (1987). *Power. A philosophical analysis*. Manchester, UK: Manchester University Press. (2nd ed., 2002)

Ramsey, F. P. (1931). *Truth and probability*. In R. B. Braithwaite (Ed.), *Foundations of mathematics and other essays*, pp. 156–198, London: Routledge and P. Kegan, 1931 [Reprinted in F. P. Ramsey, *Philosophical papers*, pp. 52–94, 1990, Cambridge: Cambridge University Press]

Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press. (Quoted from the revised edition 1999)

Uffink, J. (2011). Subjective probability and statistical physics. In C. Beisbart & S. Hartmann (Eds.), *Probabilities in physics* (pp. 25–50). Oxford: Oxford University Press.

A Probabilistic Re-View on Felsenthal and Machover's "The Measurement of Voting Power"

Olga Birkmeier and Friedrich Pukelsheim

1 The Book

Dan Felsenthal and Moshé Machover's (1998) monograph on *The Measurement of Voting Power* served a double purpose, of concisely presenting the state of the art of the theory of weighted voting systems, and of initiating novel strains of research in the area. The authors achieved these goals by a careful use of the mathematical tools, game theory and probability theory. The mathematical frame was developed not in an ivory tower seclusion, but along pertinent applications such as US-American court cases, or the Council of Ministers of the European Union. The interplay of ideal theory and concrete applications proved most fertile.

In Augsburg we repeatedly worked through the book in the course of seminars for our students who have a strong background in probability theory and statistics. Therefore we paid particular attention to the book's probabilistic language, and experimented with the technical vocabulary, in order to optimize communication with non-mathematical contemporaries. An instant stumbling stone was felt to be the phonetic closeness of two central notions of the subject, *voting weight* and *voting power*. In German they translate into *Stimmgewicht* and *Stimmkraft*. Since the German language puts a strong emphasis on the first syllable of a compound word, a negligent speaker may offer the audience an audible *Stimm*..., followed by a murmured ...*something*, thus completely missing the point. For this reason we tried to separate the notions more clearly. We kept *voting weight*, but replaced *absolute voting power* by *influence probability*, and *relative voting power* by *power share*. The term *share* indicates that the ensemble of these indices totals unity, whence they form a *power distribution*.

O. Birkmeier • F. Pukelsheim (✉)
Institut für Mathematik, Universität Augsburg, D-86315 Augsburg, Germany
e-mail: Pukelsheim@Math.Uni-Augsburg.De

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare, DOI 10.1007/978-3-319-05158-1_7,

In the present paper we explicate the probabilistic approach, in as far as we found it telling and conducive. The approach is by no means new. It dates back at least to Straffin (1978, 1988), and the Felsenthal and Machover (1998) monograph makes excellent use of it. Nevertheless we believe that a 're-view' on its role may prove useful.

In their final Chap. 8, Felsenthal and Machover (1998) make a point to take abstentions seriously. We maintain that ternary voting profiles provide a sufficiently general reference set supporting both, ternary decision rules that permit abstentions, and binary decision rules that are restricted to Yea-Nay voting (Sect. 2). The ensuing development depends on the probability distribution adopted. A model is truly ternary when it assigns positive weights to voting profiles with at least one abstention. The case of abstention probability zero leads back to the binary setting. The Penrose/Banzhaf models (Sect. 3) and the Shapley/Shubik models (Sect. 4) come with abstention probabilities $t \in [0, 1)$ that afford a smooth transition between ternary and binary settings. We conclude with an outlook on bloc decision rules, a prime example being provided by the Council of Ministers of the European Union (Sect. 5).

2 Ternary Voting Profiles

Let N denote an *assembly* consisting of finitely many *agents* j. When a proposal is tabled and a vote is taken, the results are recorded as a vector $a = (a_j)_{j \in N}$, a *voting profile*. The vote of agent j is reported as $a_j = \textit{yea}$ when j votes Yea, $a_j = \textit{nay}$ when j votes Nay, or $a_j = \textit{abstain}$ when j abstains. The natural ordering among these values is $\textit{nay} \le \textit{abstain} \le \textit{yea}$. Felsenthal and Machover (1998, p. 282) use the coding $\textit{nay} = -1$, $\textit{abstain} = 0$, and $\textit{yea} = 1$.

The ensemble of all voting profiles constitutes the *ternary profile space*

$$\Omega_N = \left\{ (a_j)_{j \in N} \;\middle|\; a_j \in \{\textit{nay}, \textit{abstain}, \textit{yea}\}, \text{ for all } j \in N \right\}.$$

Every profile $a \in \Omega_N$ induces a *region of growing acceptance* consisting of those profiles b that express at least as much acceptance as is reported in a,

$$[a, \textit{yea}] = \left\{ (b_j)_{j \in N} \in \Omega_N \;\middle|\; a_j \le b_j, \text{ for all } j \in N \right\}.$$

A subset $W_N \subseteq \Omega_N$ is called a *decision rule* when it satisfies the three properties

$$[a, \textit{yea}] \subseteq W_N, \text{ for all } a \in W_N, \tag{1}$$

$$(\textit{yea}, \ldots, \textit{yea}) \in W_N, \tag{2}$$

$$(\textit{abstain}, \ldots, \textit{abstain}) \notin W_N. \tag{3}$$

The profiles a that constitute the subset W_N are called *winning*, in the sense that a proposal is carried if and only if a belongs to W_N. We do not consider systems in which the final outcome might be a tie. Therefore the complement $W_N^C = \Omega_N \setminus W_N$ comprises the profiles that are *loosing*. Thus a subset W_N is a decision rule if and only if (1) it is *acceptance-monotonic*: if a is winning and b reports at least as much acceptance as does a, then b is also winning, (2) unanimous acceptance is winning, and (3) unanimous abstention is loosing.

Now we fix some decision rule W_N, and investigate its merits from the point of view of agent j. Two events transpire to be of particular interest. First, there is the set $A_j(W_N)$ of *agreeable profiles*, when j agrees with the final outcome. Second, there is the set $C_j(W_N)$ of *critical profiles*, when the vote of j is decisive to turn the profile winning or loosing. Let the notation $(a_i)_{i \neq j} \,\&\, (yea)_j$ represent the profile where the votes of the other agents $i \neq j$ are concatenated with a Yea from agent j. Similarly $(a_i)_{i \neq j} \,\&\, (nay)_j$ is to indicate that the votes of the others is completed with j's Nay. The two events mentioned may then be described as follows:

$$A_j(W_N) = \left\{ a \in W_N \;\middle|\; a_j = yea \right\} \cup \left\{ a \in W_N^C \;\middle|\; a_j = nay \right\},$$

$$C_j(W_N) = \left\{ a \in \Omega_N \;\middle|\; (a_i)_{i \neq j} \,\&\, (yea)_j \in W_N \text{ and } (a_i)_{i \neq j} \,\&\, (nay)_j \in W_N^C \right\}.$$

So far the exposition is descriptive and qualitative. It is only now that we consider quantitative indices. All of them originate from a probability measure P given on the ternary profile space Ω_N, with some of them being peculiar to an agent j:

$P[W_N]$	the *efficiency* of the decision rule W_N,
$P\big[A_j(W_N)\big]$	the *success probability* of agent j,
$P\big[C_j(W_N)\big]$	the *influence probability* of agent j,
$P\big[C_j(W_N)\big] / \Sigma_P(W_N)$	the *power share* of agent j, utilizing
$\Sigma_P(W_N) = \sum_{i \in N} P\big[C_i(W_N)\big]$	the *influence sensitivity* of the decision rule W_N.

The indices coincide with those in the monograph (Felsenthal and Machover 1998), except that in The Book they are related to specific distributions, namely the Penrose/Banzhaf and Shapley/Shubik distributions in their variants with abstention probability equal to zero (see below). In particular, our notion of Penrose/Banzhaf influence probability of agent j is the same as their *Banzhaf power* (or absolute Banzhaf index) of j, and our power share of agent j coincides with their *Banzhaf index* of voting power (or relative Banzhaf index) of j. Our motivation for not specializing the probabilistic assumptions too early is that there are results like Theorem 1 below which hold quite generally. To this end we need to introduce some notation.

The *dual profile* dual(a) of a ternary voting profile $a \in \Omega_N$ is defined by reversing the votes of all agents $j \in N$,

$$\big(\text{dual}(a)\big)_j = \begin{cases} \textit{nay} & \text{in case } a_j = \textit{yea}, \\ \textit{abstain} & \text{in case} a_j = \textit{abstain}, \\ \textit{yea} & \text{in case} a_j = \textit{nay}. \end{cases}$$

A distribution P is said to be *selfdual* when a voting profile and its dual are assigned identical probabilities, $P[\{a\}] = P\big[\{\text{dual}(a)\}\big]$.

A distribution P is said to be *exchangeable* when it remains invariant under all permutations of the assembly N. In the presence of exchangeability, a maximal invariant statistic tallies the yeas, nays, and abstentions of a voting profile $a \in \Omega_N$ into the three counts Yea(a), Nay(a), and Abst(a), respectively.

The *success margin* $\sigma(W_N)(a)$ is defined to be the difference between the number of those who vote in favor of the final outcome, and those who vote against it,

$$\sigma(W_N)(a) = \begin{cases} \text{Yea}(a) - \text{Nay}(a) & \text{in case } a \in W_N, \\ \text{Nay}(a) - \text{Yea}(a) & \text{in case } a \in W_N^C. \end{cases}$$

Two decision rules deserve special attention. The first is the *unanimity rule* U_N, signaling acceptance when nobody is objecting, and the second is the *straight majority rule* M_N, requiring the Yeas to outnumber the Nays,

$$U_N = \Big\{(a_j)_{j\in N} \in \Omega_N \;\Big|\; \text{Yea}(a) > 0 = \text{Nay}(a)\Big\},$$

$$M_N = \Big\{(a_j)_{j\in N} \in \Omega_N \;\Big|\; \text{Yea}(a) > \text{Nay}(a)\Big\}.$$

Theorem 1. *Let the ternary profile space Ω_N be equipped with be a selfdual and exchangeable probability distribution P.*

Then every decision rule W_N has its expected success margin lying between the expected success margins of the unanimity rule and of the straight majority rule,

$$\mathrm{E}_P\big[\sigma(U_N)\big] \leq \mathrm{E}_P\big[\sigma(W_N)\big] \leq \mathrm{E}_P\big[\sigma(M_N)\big].$$

Proof. See Proposition 4.1 in Birkmeier et al. (2011). □

The unanimity rule and the straight majority rule are two instances of the wider class of *weighted decision rules* $W_N[q; (w_j)_{j\in N}]$. Such a rule is determined by a *quota* $q \in [0, 1)$, and *voting weights* $w_j > 0$ for agents $j \in N$. For a given voting profile a let $\text{YCW}(a) = \sum_{j\in N: a_j = yea} w_j$ designate the *Yea-voters' cumulative weight*, and $\text{NCW}(a) = \sum_{j\in N: a_j = nay} w_j$ the *Nay-voters' cumulative weight*. A ternary voting profile a is defined to be winning, $a \in W_N[q; (w_j)_{j\in N}]$, when the Yea-voters' cumulative weight exceeds the fraction q of the cumulative weight of all non-abstainers, $\text{YCW}(a) > q \cdot \big(\text{YCW}(a) + \text{NCW}(a)\big)$.

3 The Penrose/Banzhaf Model

The *Penrose/Banzhaf distribution* P_N^t assumes that all agents act independently, abstain with a common *abstention probability* $t \in [0, 1)$, and divide the remaining likelihood $1 - t$ equally between a Yea and a Nay. In this model, a ternary voting profile $a \in \Omega_N$ carries the probability

$$P_N^t[\{a\}] = \frac{1}{2^{\mathrm{Yea}(a)+\mathrm{Nay}(a)}} (1-t)^{\mathrm{Yea}(a)+\mathrm{Nay}(a)} \, t^{\mathrm{Abst}(a)}.$$

When the ternary parameter t vanishes, voting profiles that contain an abstention are assigned zero probability. Thus a profile carries positive mass only when every agent votes Yea or Nay. That is, with $t = 0$ the ternary Penrose/Banzhaf model reduces to the familiar binary Penrose/Banzhaf model. The ternary Penrose/Banzhaf model thus embraces the binary Penrose/Banzhaf model as a degenerate case. A sample result is provided by Theorem 2.

Theorem 2. *Let the ternary profile space Ω_N be equipped with the Penrose/Banzhaf distribution P_N^t, with abstention probability $t \in [0, 1)$, and let W_N be an arbitrary decision rule.*

(i) *For all agents $j \in N$, success and influence probabilities are related through*

$$P_N^t\big[A_j(W_N)\big] = \frac{1-t}{2} + \frac{1-t}{2} P_N^t\big[C_j(W_N)\big].$$

(ii) *The influence sensitivity and the expected success margin of W_N fulfill*

$$\Sigma_{P_N^t}(W_N) = \frac{1}{1-t} \mathrm{E}_{P_N^t}\big[\sigma(W_N)\big].$$

Proof. See Propositions 5.1 and 5.2 in Birkmeier et al. (2011). □

With $t = 0$, this coincides with the results in Theorems 3.2.16 and 3.3.5 in Felsenthal and Machover (1998), see also Ruff and Pukelsheim (2010). Birkmeier (2011, Satz 2.3.4) presents a version of part (i) dealing with a slightly larger set of profiles that are considered a success for agent j, namely those that are agreeable to agent j combined with those wherein j abstains (and which might be considered "weakly agreeable").

4 The Shapley/Shubik Model

The *Shapley/Shubik distribution* S_N^t on Ω_N is built up in three stages. The first stage, dealing with abstentions, is new. We propose to assume all agents to abstain independently, with a common abstention probability $t \in [0, 1)$. Under this assumption the number ℓ of those who abstain follows a binomial distribution, $\frac{n!}{\ell!(n-\ell)!}t^{\ell}(1-t)^{n-\ell}$. The second and third stages are standard. The number of Yea-voters k is taken to attain each of its possible values $0,\dots,n-\ell$ with the same probability, $1/\big(n-\ell+1\big)$. Third, each of the $\frac{n!}{k!\ell!(n-k-\ell)!}$ profiles with k Yeas, ℓ abstentions, and $n-k-\ell$ Nays is considered equally likely. Thus, with some of the factorial terms canceling out and after re-substituting Yea(a) for k and Yea(a) + Nay(a) for $n-\ell$, the total probability of a ternary voting profile $a \in \Omega_N$ becomes

$$S_N^t[\{a\}] = \frac{\mathrm{Yea}(a)!\,\mathrm{Nay}(a)!}{\big(\mathrm{Yea}(a)+\mathrm{Nay}(a)+1\big)!}\,(1-t)^{\mathrm{Yea}(a)+\mathrm{Nay}(a)}\,t^{\mathrm{Abst}(a)}.$$

In binary models, it is well-known that every decision rule W_N has Shapley/Shubik influence sensitivity equal to unity. This entails two intriguing consequences, that the Shapley/Shubik sensitivity is insensitive to the specific decision rule W_N, and that the Shapley/Shubik influence probability of an agent j coincides with her or his power share. In ternary Shapley/Shubik models, the first conclusion persists, the second does not.

Theorem 3. *Let the ternary profile space Ω_N be equipped with the Shapley/Shubik distribution S_N^t, with abstention probability $t \in [0,1)$, and let n be the cardinality of the assembly N.*

Then all decision rules W_N share an identical influence sensitivity,

$$\Sigma_{S_N^t}(W_N) = \frac{1-t^n}{1-t}.$$

Proof. See Satz 2.3.9 in Birkmeier (2011). □

The right hand side is the same as $1+t+\cdots+t^{n-1}$. Hence its limit equals n, the number of agents, as the abstention probability t tends to unity. This is quite plausible since, with the likelihood of abstention growing, each of the n agents is getting to be more and more critical when casting a clear Yea- or Nay-vote, and in the end acquires an influence probability equal to unity.

5 The EU Council of Ministers

In some applications the grand assembly N is partitioned into disjoint subsets, called *blocs*. The associated two-tier voting system is composed of internal decision rules within blocs, and a second-level decision rule among bloc delegates. An example is the Council of Ministers of the European Union, where the entirety of the Union citizens, N, is partitioned into the 27 blocs of its Member States' citizenries that are represented by their Ministers.

In the binary Penrose/Banzhaf bloc model, the influence probability of citizen $j \in N$ in a two-tier system typically factorizes into the product of the internal influence probability of j in her or his bloc B, times the second-level influence probability of bloc B relative to the specified bloc partitioning, see Straffin (1978), Felsenthal and Machover (2002), Laruelle and Valenciano (2004), or Ruff and Pukelsheim (2010). These product formulas generalize to carry over to ternary Penrose/Banzhaf bloc models, see Birkmeier (2011) for a *bottom up* construction (Satz 5.1.3) as well as for a *top down* construction (Satz 5.1.6).

The analysis may be employed to design optimal decision rules that permit abstentions. The underlying notion of optimality builds on a weighted average of the diplomatic *one state, one vote* principle that underlies international relations among Member States, and of the democratic *one person, one vote* principle that would apply to the Union citizens, see Laruelle and Widgrén (1998) and Satz 5.3.3 in Birkmeier (2011). However, it is by no means evident whether the Treaty of Lisbon (2010) would support the two equality principles and, if so, whether they may be mixed into a single optimality criterion.

Nevertheless, a statistical evaluation of previous decision rules used in the EU Council of Ministers leads to the estimates reported in Sect. 1 of Birkmeier (2011). They suggest that, in the past, the Union functioned with a mixture that puts a weight of 10 % on the diplomatic equality principle, and a complementary 90 % weight on the democratic equality principle. With these weightings, the optimal quota is found to be 60.98 %, see Birkmeier (2011, p. 117). This is slightly below the quota of 61.6 % proposed in the Jagiellonian Compromise of Słomczyński and Życzkowski (2010).

The mixture criterion is roughly in line with the composition of the European Parliament where each Member State is guaranteed six seats out of a total of 751 seats. That is, 20 % of the seats are preassigned to the Member States obeying the diplomatic equality principle of *one state, one vote*. The remaining 80 % then might be allocated via a proportional representation apportionment method to honor the democratic equality principle of *one person, one vote*, as proposed in the Cambridge Compromise of Grimmett et al. (2011).

References

Birkmeier, O. (2011). *Machtindizes und Fairness-Kriterien in gewichteten Abstimmungssytemen mit Enthaltungen*. Augsburger Schriften zur Mathematik, Physik und Informatik, Band 18. Berlin: Logos.

Birkmeier, O., Käufl, A., & Pukelsheim, F. (2011). Abstentions in the German Bundesrat and ternary decision rules in weighted voting systems. *Statistics & Decisions*, *28*, 1–16.

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

Felsenthal, D.S., & Machover, M. (2002). Annexations and alliances: When are blocs advantageous a priori? *Social Choice and Welfare*, *19*, 295–312.

Grimmett, G., Laslier, J. -F., Pukelsheim, F., Ramírez-González, V., Rose, R., Słomczyński, W., Zachariasen, M., Zyczkowski, K. (2011). *The allocation between the EU Member States of the seats in the European Parliament*. European Parliament, Directorate-General for Internal Policies, Policy Department C: Citizen's Rights and Constitutional Affairs, Note 23.03.2011 (PE 432.760).

Laruelle, A., & Valenciano, F. (2004). On the meaning of Owen-Banzhaf coalitional value in voting situation. *Theory and Decision*, *56*, 113–123.

Laruelle, A., & Widgrén, M. (1998). Is the allocation of voting power among the EU States fair? *Public Choice*, *94*, 317–339.

Ruff, O., & Pukelsheim, F. (2010). A probabilistic synopsis of binary decision rules. *Social Choice and Welfare*, *35*, 501–516.

Słomczyński, W., & Życzkowski, K. (2010). Jagiellonian compromise – an alternative voting system for the Council of the European Union. In M. A. Cichocki, & Życzkowski, K. (Eds.), *Distribution of power and voting procedures in the European Union* (pp. 43–58). London: Ashgate.

Straffin, P. D. (1978). Probability models for power indices. In P. C. Ordeshook (Ed.), *Game theory and political science* (pp. 477–510). New York: New York University Press.

Straffin, P. D. (1988). The Shapley-Shubik and Banzhaf power indices as probabilities. In A. E. Roth (Ed.), *The Shapley value – essays in honor of Lloyd S. Shapley* (pp. 71–81). Cambridge: Cambridge University Press.

Treaty of Lisbon. (2010). Treaty on European Union, Articles 1–19. *Official Journal of the European Union*, *C83*, (30.2.2010) 13–27.

Part III
Power in Two-Tier Voting Systems

Square Root Voting System, Optimal Threshold and π

Karol Życzkowski and Wojciech Słomczyński

1 Introduction

Recent political debate on the voting system used in the Council of Ministers of the European Union stimulated research in the theory of indirect voting, see e.g. Felsenthal and Machover (2001), Leech (2002), Andjiga et al. (2003), Pajala and Widgrén (2004), and Beisbart et al. (2005). The *double majority* voting system, adopted for the Council by The Treaty of Lisbon in December 2007 is based on two criteria: "per capita" and "per state". This system apparently reflects the principles of equality of Member States and that of equality of citizens. However, as recently analyzed by various authors Baldwin and Widgrén (2004), Ade (2006), Słomczyński and Życzkowski (2006), Algaba et al. (2007), Hosli (2008), Bârsan-Pipu and Tache (2009), Kirsch (2010), Moberg (2010), Leech and Aziz (2010), Pukelsheim (2010), and Słomczyński and Życzkowski (2010), in such a system the large states gain

---

This paper was presented to the *Voting Power in Practice Symposium* at the London School of Economics, 20–22 March 2011, sponsored by the Leverhulme Trust.

K. Życzkowski (✉)
Institute of Physics, Jagiellonian University, ul. Reymonta 4, 30-059 Kraków, Poland

Center for Theoretical Physics, Polish Academy of Sciences, Al. Lotników 32/46, 02-668 Warszawa, Poland
e-mail: karol@tatry.if.uj.edu.pl

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

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare, DOI 10.1007/978-3-319-05158-1_8,

a lot of power from the direct link to population, while the smallest states derive disproportionate power from the other criterion. The combined effect saps influence away from all medium-sized countries. Ironically, a similar conclusion follows from a book by Lionel Penrose, who wrote already in 1952 (Penrose 1952):

> If two votings were required for every decision, one on a per capita basis and the other upon the basis of a single vote for each country, this system would be inaccurate in that it would tend to favor large countries.

To quantify the notion of voting power, mathematicians introduced the concept of power index of a member of the voting body, which measures the probability that his vote will be decisive in a hypothetical ballot: Should this member decide to change his vote, the winning coalition would fail to satisfy the qualified majority condition. Without any further information about the voting body it is natural to assume that all potential coalitions are equally likely. This very assumption leads to the concept of Banzhaf(-Penrose) index called so after John Banzhaf, an American attorney, who introduced this index independently in 1965 (Banzhaf 1965).

Note that this approach is purely normative, not descriptive: we are interested in the potential voting power arising from the voting procedure itself. Calculation of the voting power based on the counting of majority coalitions is applicable while analyzing institutions in which alliances are not permanent, but change depending upon the nature of the matter under consideration.

To design a representative voting system, i.e. the system based on the democratic principle, that the vote of any citizen of any Member State is of equal worth, one needs to use a weighted voting system. Consider elections of the government in a state with population of size N. It is easy to imagine that an average German citizen has smaller influence on the election of his government than, for example, a citizen of the neighboring Luxembourg. Analyzing this problem in the context of voting in the United Nations just after the World War II Penrose showed, under some natural assumptions, that in such elections the voting power of a single citizen decays as one over square root of N. Thus, the system of indirect voting applied to the Council is representative, if the voting power of each country is proportional to the square root of N, so that both factors cancel out. This statement is known in the literature under the name of the *Penrose square root law* (Penrose 1946; Felsenthal and Machover 1998). It implies that the voting power of each member of the EU Council should behave as $\sqrt{N}$ and such voting systems have been analyzed in this context by several experts since late 1990s (Felsenthal and Machover 1997; Laruelle and Widgrén 1998).

It is challenging to explain this fact in a way accessible to a wide audience (Życzkowski et al. 2006; Kirsch et al. 2007; Pukelsheim 2007; Pöppe 2007). A slightly paradoxical nonlinearity in the result of Penrose is due to the fact that voting in the Council should be considered as a two-tier voting system: Each member state elects a government, which delegates its representative to the Council. Any representative has to say *"Yes"* or *"No"* on behalf of his state in every voting organized in the Council. The key point is that in such a voting each member of the Council cannot split his vote. Making an idealistic assumption that the vote of a

Minster in the Council represents the will of the majority of the citizens of the state he represents, his vote "*Yes*" means only that a majority of the population of his state supports this decision, but does not reflect the presence of a minority.

Consider an exemplary issue to be voted in the Council and assume that the preferences of the voters in each state are known. Assume hypothetically that a majority of population of Malta says "*Yes*" on a certain issue, the votes in Italy split as 30 millions "*Yes*" and 29 millions "*No*", while all 43 millions of citizens of Spain say "*No*". A member of the Council from Malta follows the will of the majority in his state and votes "*Yes*". So does the representative of Italy. According to the double majority voting system his vote is counted on behalf of the total number of 59 millions of the population of Italy. Thus these voting rules allow 30 millions of voters in Italy to over-vote not only the minority of 29 millions in their state (which is fine), but also, with the help of less than half a million of people from Malta, to over-vote 43 millions of Spaniards.

This pedagogical example allows one to conclude that the double majority voting system would work perfectly, if all voters in each member state had the same opinion on every issue. Obviously such an assumption is not realistic, especially in the case of the European states, in which the citizens can nowadays afford the luxury of an independent point of view. In general, if a member of the Council votes "*Yes*" on a certain issue, in an ideal case one may assume that the number of the citizens of his state which support this decision varies from 50 % till 100 % of the total population. In practice, no concrete numbers for each state are known, so to estimate the total number of European citizens supporting a given decision of the Council one has to rely on statistical reasoning.

To construct the voting system in the Council with voting powers proportional to the square root of populations one can consider the situation, where voting weights are proportional to the square root of populations and the Council takes its decision according to the principle of a *qualified majority*. In other words, the voting in the Council yields acceptance, if the sum of the voting weights of all Ministers voting "*Yes*" exceeds a fixed quota q, set for the qualified majority. From this perspective the quota q can be treated as a free parameter (Leech and Machover 2003; Machover 2010), which may be optimized in such a way that the mean discrepancy Δ between the voting power (measured by the Banzhaf index) and the voting weight of each member state is minimal.

In the case of the population in the EU consisting of 25 member states it was shown (Słomczyński and Życzkowski 2004; Życzkowski et al. 2006) that the value of the optimal quota q_* for qualified majority in the Penrose's square root system is equal to 62.0 %, while for EU-27 this number drops down to 61.5 % (Słomczyński and Życzkowski 2006, 2007). Furthermore, the optimal quota can be called *critical*, since in this case the mean discrepancy $\Delta(q_*)$ is very close to zero and thus the voting power of every citizen in each member state of the Union is practically equal. This simple scheme of voting in the EU Council based on the square root law of Penrose supplemented by a rule setting the optimal quota to q_* happens to give larger voting powers to the largest EU than the Treaty of Nice, but smaller ones than the Treaty of Lisbon. Therefore this voting system has been dubbed by the media as the *Jagiellonian Compromise*.

It is known that the existence of the critical quota q_*, is not restricted to the particular distribution of the population in the European Union, but it is also characteristic of a generic distribution of the population (Słomczyński and Życzkowski 2004, 2006; Chang et al. 2006). The value of the critical quota depends on the particular distribution of the population in the "union", but even more importantly, it varies considerably with the number M of member states. An explicit approximate formula for the critical quota was derived in Słomczyński and Życzkowski (2007). It is valid in the case of a relatively large number of the members of the "union" and in the asymptotic limit, $M \to \infty$, the critical quota tends to 50 %, in consistence with the so-called *Penrose limit theorem* (Lindner and Machover 2004).

On one hand it is straightforward to apply this explicit formula for the current population of all member states of the existing European Union, as well as to take into account various possible scenarios of a possible extension of the Union. On the other hand, if the number of member states is fixed, while their populations vary in time, continuous update of the optimal value for the qualified majority may be cumbersome and unpractical. Hence one may try to neglect the dependence on the particular distribution of the population by selecting for the quota the mean value of $\langle q \rangle$, where the average is taken over a sample of random population distributions, distributed uniformly in the allowed space of M-point probability distributions. In this work we perform such a task and derive an explicit, though approximate, formula for the average critical quota.

This paper is organized as follows. In Sect. 2 devoted to the one-tier voting system, we recall the definition of Banzhaf index and review the Penrose square root law. In Sect. 3, which concerns the two-tier voting systems, we describe the square root voting system and analyze the average number of misrepresented voters. Section 4 is devoted to the problem of finding the optimal quota for the qualified majority. It contains the key result of this paper: derivation of a simple approximate formula for the average optimal quota, which depends only on the number M of the member states and is obtained by averaging over an ensemble of random distributions of the population of the "union".

2 One Tier Voting

Consider a voting body consisting of M voters voting according to the qualified majority rule. Assume that the weights of the votes need not to be equal, which is typical e.g. in the case of an assembly of stockholders of a company: the weight of the vote of a stockholder depends on the number of shares he or she possesses. It is worth to stress that, generally, the voting weights do not directly give the voting power.

To quantify the a priori voting power of any member of a given voting body game theorists introduced the notion of a power index. It measures the probability that a member's vote will be decisive in a hypothetical ballot: should this player decide to change its vote, the winning coalition would fail to satisfy the qualified majority condition. In the game theory approach to voting such a player is called *pivotal*.

The assumption that all potential coalitions of voters are equally likely leads to the concept of the Banzhaf index (Penrose 1946; Banzhaf 1965). To compute this power index for a concrete case one needs to enumerate all possible coalitions, identify all winning coalitions, and for each player find the number of cases in which his vote is decisive.

Let M denote the number of voters and ω the total number of all winning coalitions, that satisfy the qualified majority condition. Assume that ω_k denotes the number of winning coalitions that include the k-th player; where $k = 1, \ldots, M$. Then the Banzhaf index of the k-th voter reads

$$\psi_k \;:=\; \frac{\omega_k - (\omega - \omega_k)}{2^{M-1}} = \frac{2\omega_k - \omega}{2^{M-1}} \;. \tag{1}$$

To compare these indices for decision bodies consisting of different number of players, it is convenient to define the *normalized Banzhaf (-Penrose) index*:

$$\beta_k \;:=\; \frac{\psi_k}{\sum_{i=1}^{M} \psi_i} \tag{2}$$

such that $\sum_{i=1}^{M} \beta_i = 1$.

In the case of a small voting body such a calculation is straightforward, while for a larger number of voters one has to use a suitable computer program.

2.1 Square Root Law of Penrose

Consider now the case of N members of the voting body, each given a single vote. Assume that the body votes according to the standard majority rule. On one hand, since the weights of each voter are equal, so must be their voting powers. On the other hand, we may ask, what happens if the size N of the voting body changes, for instance, if the number of eligible voters gets doubled, how does this fact influence the voting power of each voter?

For simplicity assume for a while that the number of voters is odd, $N = 2j + 1$. Following original arguments of Penrose we conclude that a given voter will be able to effectively influence the outcome of the voting only if the votes split half and half: If the vote of j players would be "*Yes*" while the remaining j players vote "*No*", the role of the voter we analyze will be decisive.

Basing upon the assumption that all coalitions are equally likely one can ask, how often such a case will occur? In mathematical language the model in which this assumption is satisfied is equivalent to the *Bernoulli scheme*. The probability that out of $2j$ independent trials we obtain k successes reads

$$P_k := \binom{2j}{k} p^k (1-p)^{2j-k} \;, \tag{3}$$

where p denotes the probability of success in each event. In the simplest symmetric case we set $p = 1 - p = 1/2$ and obtain

$$P_j = \left(\frac{1}{2}\right)^{2j} \frac{(2j)!}{(j!)^2}. \tag{4}$$

For large N we may use the Stirling approximation for the factorial and obtain the probability ψ that the vote of a given voter is decisive

$$\psi = P_j \sim 2^{-2j} \frac{(2j/e)^{2j} \sqrt{4\pi j}}{[(j/e)^j \sqrt{2\pi j}]^2} = \frac{1}{\sqrt{\pi j}} \sim \sqrt{\frac{2}{\pi N}}. \tag{5}$$

For N even we get the same approximation. In this way one can show that the voting power of any member of the voting body depends on its size as $1/\sqrt{N}$, which is the *Penrose square root law*. The above result is obtained under the assumption that the votes of all citizens are uncorrelated. A sound mathematical investigation of the influence of possible correlations between the voting behavior of individual citizens for their voting power has been recently presented by Kirsch (2007). It is easy to see that due to strong correlations certain deviations from the square root law have to occur, since in the limiting case of unanimous voting in each state (perfect correlations), the voting power of a single citizen from a state with population N will be inversely proportional to N.

The issue that the assumptions leading to the Penrose law are not exactly satisfied in reality was raised many times in the literature, see, e.g. Gelman et al. (2002, 2004), also in the context of the voting in the Council of the European Union (Laruelle and Valenciano 2008). However, it seems not to be easy to design a rival model voting system which correctly takes into account the essential correlations, varying from case to case and evolving in time. Furthermore, it was argued (Kirsch 2007) that the strength of the correlations between the voters tend to decrease in time. Thus, if one is to design a voting system to be used in the future in the Council of the European Union, it is reasonable to consider the idealistic case of no correlations between individual voters. We will follow this strategy and in the sequel rely on the square root law of Penrose.

2.2 Pivotal Voter and the Return Probability in a Random Walk

It is worth to emphasize that the square root function appearing in the above derivation is typical to several other reasonings in mathematics, statistics and physics. For instance, in the analyzed case of a large voting body, the probability distribution P_k in the Bernoulli scheme can be approximated by the Gaussian distribution with the standard deviation being proportional to $1/\sqrt{N}$. It is also instructive to compare the above voting problem with a simple model of a random walk on the one dimensional lattice.

Assume that a particle subject to external influences in each step jumps a unit distance left or right with probability one half. What is the probability that it returns to the initial position after N steps? It is easy to see that the probability scales as $1/\sqrt{N}$, since the answer is provided by exactly the same reasoning as for the Penrose law.

Consider an ensemble of particles localized initially at the zero point and performing such a random walk on the lattice. If the position of a particle at time n differs from zero, in half of all cases it will jump towards zero, while in the remaining half of cases it will move in the opposite direction. Hence the *mean* distance $\langle D \rangle$ of the particle from zero will not change. On the other hand, if at time n the particle happened to return to the initial position, in the next step it would certainly jump away from it, so the mean distance from zero would increase by one.

To compute the mean distance form zero for an ensemble of random particles performing N steps, we need to sum over all the cases, when the particle returns to the initial point. Making use of the previous result, that the return probability $P(n)$ at time n behaves as $1/\sqrt{n}$, we infer that during the time N the mean distance behaves as

$$\langle D(N) \rangle \approx \sum_{n=1}^{N} P(n) \approx \sum_{n=1}^{N} \frac{1}{\sqrt{n}} \sim \sqrt{N}. \tag{6}$$

This is just one formulation of the *diffusion law*. As shown, the square root of Penrose is closely related with some well known results from mathematics and physics, including the Gaussian approximation of binomial distribution and the diffusion law.

3 Two Tier Voting

In a two-tier voting system each voter has the right to elect his representative, who votes on his behalf in the upper chamber. The key assumption is that, on one hand, he should represent the will of the population of his state as best he can, but, on the other hand, he is obliged to vote "*Yes*" or "*No*" in each ballot and cannot split his vote. This is just the case of voting in the Council of the EU, since citizens in each member state choose their government, which sends its Minister to represent the entire state in the Council.

These days one uses in the Council the triple majority system adopted in 2001 in the *Treaty of Nice*. The Treaty assigned to each state a certain number of "weights", distributed in an ad hoc fashion. The decision of the Council is taken if the coalition voting in favour of it satisfies three conditions:

(a) it is formed by the standard majority of the member states,
(b) states forming the coalition represent more then 62 % of the entire population of the Union,
(c) the total number of weights of the "*Yes*" votes exceeds a quota equal to approximately 73.9 % of all weights.

Although all three requirements have to be fulfilled simultaneously, detailed analysis shows that condition (c) plays a decisive role in this case: if it is satisfied, the two others will be satisfied with a great likelihood as well (Felsenthal and Machover 2001; Leech 2002).

Therefore, the voting weights in the Nice system play a crucial role. However, the experts agree (Felsenthal and Machover 2001; Pajala and Widgrén 2004) that the choice of the weights adopted is far from being optimal. For instance the voting power of some states (including e.g. Germany and Romania) is significantly smaller than in the square root system. This observation is consistent with the fact that Germany was directly interested to abandon the Nice system and push toward another solution that would shift the balance of power in favor of the largest states.

In the double majority voting system, adopted in December 2007 in Lisbon, one gave up the voting weights used to specify the requirement (c) and decided to preserve the remaining two conditions with modified majority quotas. A coalition is winning if:

(a') it is formed by at least 55 % of the members states,
(b') it represents at least 65 % of the population of the Union.

Additionally, every coalition consisting of all but three (or less) countries is winning even if it represents less than 65 % of the population of the Union.

The double majority system will be used in the Council starting from the year 2014. However, a detailed analysis by Moberg (2010) shows that in this concrete case the "double majority" system is not really double, as the per capita criterion (b') plays the dominant role here. In comparison with the Treaty of Nice, the voting power index will increase for the four largest states of the Union (Germany, France, the United Kingdom and Italy) and also for the smallest states. To understand this effect we shall analyze the voting system in which the voting weight of a given state is directly proportional to its population.

3.1 Voting Systems with Per Capita Criterion

The idea "one citizen—one vote" looks so natural and appealing, that in several political debates one often did not care to analyze in detail its assumptions and all its consequences. It is somehow obvious that a minister representing a larger (if population is considered) state should have a larger weight during each voting in the EU Council. On the other hand, one needs to examine whether the voting weights of a minister in the Council should be proportional to the population he represents. It is clear that this would be very much the case, if one could assume that all citizens in each member state share the very same opinion in each case.

However, this assumption is obviously false, and nowadays we enjoy in Europe the freedom to express various opinions on every issue. Let us then formulate the question, how many citizens from his state each minister actually represents in an exemplary voting in the Council? Or to be more precise, how many voters from a given state with population N share in a certain case the opinion of their representative? We do not know!

Under the idealizing assumption that the minister always votes according to the will of the majority of citizens in his state, the answer can vary from $N/2$ to N. Therefore, the difference between the number of the citizens supporting the vote of their minister and the number of those who are against it can vary from 0 to N. In fact it will vary from case to case in this range, so an assumption that it is always proportional to N is false. This crucial issue, often overlooked in popular debates, causes problems with representativeness of a voting system based on the "per capita" criterion.

There is no better way to tackle the problem as to rely on certain statistical assumptions and estimate the *average* number of "satisfied citizens". As such an analysis is performed later in this paper, we shall review here various arguments showing that a system with voting weights directly proportional to the population is advantageous to the largest states of the union.

Consider first a realistic example of a union of nine states: a large state A, with 80 millions of citizens and eight small states from B to I, with 10 millions each. Assume now that in a certain case the distribution of the opinion in the entire union is exactly polarized: in each state approximately 50 % of the population support the vote "*Yes*", while the other half is against. Assume now that the government of the large state is in position to establish exactly the will of the majority of citizens in their state (say it is the vote "*Yes*") and order its minister to vote accordingly. Thus the vote of this minister in the council will then be counted as a vote of 80 millions of citizens.

On the other hand, in the remaining states the probability that the majority of citizens support "*Yes*" is close to 50 %. Hence it is most likely that the votes of the ministers from the smaller states split as 4:4. Other outcomes: 5:3, 6:2, or 7:1 are less probable, but all of them result in the majority of the representative of the large state A. The outcome 8:0 is much less likely, so if we sum the votes of all nine ministers we see that the vote of the minister from the largest state will be decisive. Hence we have shown that the voting power of all citizens of the nine small states is negligible, and the decision for this model union is practically taken by the half of its population belonging to the largest state A. Even though in this example we concentrated on the "per capita" criterion and did not take into account the other criterion, it is not difficult to come up with analogous examples which show that the largest states are privileged also in the double majority system. Similarly, the smallest states of the union benefit from the "per state" criterion.

Let us have a look at the position of the minority in large states. In the above example the minority in the 80 million state can be as large as 40 million citizens, but their opinion will not influence the outcome of the voting, independently of the polarization of opinion in the remaining eight states. Thus one may conclude that in the voting system based on the "per capita" criterion, the influence of the politicians representing the majority in a large state is enhanced at the expense of the minority in this state and the politicians representing the smaller states.

Last but not least, let us compare the maximal sizes of the minority, which can arise during any voting in an EU member state. In Luxembourg, with its population of about 400,000 people, the minority cannot exceed 200,000 citizens. On the other hand, in Germany, which is a much larger country, it is possible that the minority exceeds 41 millions of citizens, since the total population exceeds 82 millions. It is then fair to say, that, due to elections in smaller states, we know the opinion of citizens in these states with a better accuracy, than in larger members of the union. Thus, as in smaller states the number of misrepresented citizens is smaller, their votes in the EU Council should be weighted by larger weights than the vote of the largest states. This very idea is realized in the weighted voting system advocated by Penrose.

3.2 Square Root Voting System of Penrose

The Penrose system for the two-tier voting is based on the square root law reviewed in Sect. 2.1. Since the voting power of a citizen in state k with population N_k scales as $1/\sqrt{N_k}$, this factor will be compensated, if the voting power of each representative in the upper chamber will behave as $\sqrt{N_k}$. Only in this way the voting power of each citizen in every state of a union consisting of M states will be equal.

Although we know that the voting power of a minister in the Council needs not coincide with the weight of his vote, as a rough approximation let us put his weights w_k proportional to the square root of the population he represents, that is $w_k = \sqrt{N_k}/\sum_{i=1}^{M}\sqrt{N_i}$.

To see a possible impact of the change of the weights let us now return to the previous example of a union of one big state and eight small ones. As the state A is eight times as large as each of the remaining states, its weight in the Penrose system will be $w_A = \sqrt{8}w_B$. As $\sqrt{8}$ exceeds 2 and is smaller than 3, we see that accepting the Penrose system will increase the role of the minority in the large state and the voting power of all smaller states. For instance, if the large state votes "*Yes*" and the votes in the eight states split as 2:6 or 1:7 in favor for "*No*", the decision will not be taken by the council, in contrast to the simple system with one "per capita" criterion. There, we have assumed that the standard majority of weights is sufficient to form a winning coalition. If the threshold for the qualified majority is increased to 54 %, also the outcome 3:5 in favor for "*No*" in the smaller states suffices to block the decision taken in the large state.

This simple example shows that varying the quota for the qualified majority considerably influences the voting power, see also Leech and Machover (2003) and Machover (2010). The issue of the selection of the optimal quota will be analyzed in detail in the subsequent section. At this point, it is sufficient to add that in general it is possible to find such a level of the quota for which the voting power β_k of the k-th state is proportional to $\sqrt{N_k}$ and, in consequence, the Penrose law is almost exactly fulfilled (Słomczyński and Życzkowski 2004, 2006).

Applying the square root voting system of Penrose combined with the optimal quota to the problem of the Council, one obtains a fair solution, in which every citizen in each member state of the Union has the same voting power, hence the same influence on the decisions taken by the Council. In this case, the voting power of each European state measured by the Banzhaf index scales as the square root of its population. This weighted voting system happens to give a larger voting power to the largest EU states (including Germany) than the Treaty of Nice but smaller than the double majority system. On the other hand, this system is more favorable to all middle size states then the double majority, so it is fair to consider it as a compromise solution. The square root voting system of Penrose is simple (one criterion only), transparent and efficient—the probability of forming a winning coalition is reasonably high. Furthermore, as discussed later, it can be easily adopted to any possible extension of the Union.

3.3 The Second Square Root Law of Morriss

To provide an additional argument in favour of the square root weights of Penrose (Felsenthal 1999), consider a model state of N citizens, of which a certain number k support a given legislation to be voted in the council. Assume that the representative of this state knows the opinion of his people and, according to the will of the majority, he votes "*Yes*" in the council if $k \geq N/2$. Then the number of citizens satisfied with his decision is k. The number $N - k$ of disappointed citizens compensates the same number of yes-votes, so the vote of the minister should effectively represent the *difference* between them, $w = k - (N - k) = 2k - N$. By our assumption concerning the majority this number is positive, but in general the effective weight of the vote of the representative should be $w = |2k - N|$.

Assume now that the votes of any of N citizens of the state are independent, and that both decisions are equally likely, so that $p = 1 - p = 1/2$. Thus, for the statistical analysis, we can use the Bernoulli scheme (3) and estimate the weight of the vote of the minister by the average using the Stirling approximation:

$$\begin{aligned}\langle w_N \rangle &= \sum_{k=0}^{N} P_k |2k - N| = \sum_{k=0}^{N} \binom{N}{k} \frac{1}{2^N} |2k - N| \\ &= \frac{\lfloor N/2 \rfloor + 1}{2^{N-1}} \binom{N}{\lfloor N/2 \rfloor + 1} \sim \sqrt{\frac{2N}{\pi}}. \end{aligned} \tag{7}$$

Here $\lfloor x \rfloor$ denotes the largest integer not greater than x. This result provides another argument in favor of the weighted voting system of Penrose: Counting all citizens of a given state, we would attribute the weights of the representative proportionally to the population N he is supposed to represent. On the other hand, if we take into account the obvious fact that not all citizens in this state share the opinion of the government on a concrete issue and consider the average number of the majority of citizens which support his decision one should weight his vote proportionally to $\sqrt{N}$. From this fact one can deduce the *second square root law of Morriss* (Morriss 2002; Felsenthal and Machover 1998; Felsenthal 1999; Laruelle and Valenciano 2008) that states that the average number of misrepresented voters in the union is smallest if the weights are proportional to the square root of the population and quota is equal to 50 %, provided that the population of each member state is large enough. Simultaneously, in this situation, the total voting power of the union measured by the sum of the Banzhaf indices of all citizens in the union is maximal.

To illustrate the result consider a model union consisting of one large state with population of 49 millions, three medium states with 16 million each and three small with 1 million citizens. For simplicity assume that the double majority system and the Penrose system are based on the standard majority of 50 %. If the polarization of opinion in each state on a given issue is as in the table below, only 39 % of the population of the union is in favor of the legislative. However, under the rules of the double majority system the decision is taken (against the will of the vast majority!), what is not the case in the Penrose system, for which the coalition gains only 10 votes out of 22, so it fails to gather the required quota.

To qualitatively understand this result, consider the minister representing the largest country G with a population of 49 millions. In the double majority system he uses his 49 votes against the will of 24 millions of inhabitants. By contrast, the minister of the small state A will misrepresent at most one half of the million of his compatriots. In other words, the precision in determining the will of all the citizens is largest in the smaller states, so the vote of their ministers should gain a higher weight than proportional to population, which is the case in the Penrose system – see Table 1.

Table 1 Case study: voting in the council of a model union of seven members under a hypothetical distribution of population and voting preferences

State	A	B	C	D	E	F	G	Total	
Population [M]	1	1	1	16	16	16	49	100	
Votes: Yes [M]	$2/3$	$2/3$	$2/3$	4	4	4	25	39	
Votes: No [M]	$1/3$	$1/3$	$1/3$	12	12	12	24	61	
State votes	1	1	1	0	0	0	1	$4/7$	Y
Minister's votes	1	1	1	0	0	0	49	$52/100$	Y
Square root weights	1	1	1	4	4	4	7	22	
Square root votes	1	1	1	0	0	0	7	$10/22$	N

Although 61 % of the total population of the union is against a legislative it will be taken by the council, if the rules of the double majority are used. The outcome of the voting according to the weighted voting system of Penrose correctly reflects the will of the majority in the union

4 Optimal Quota for Qualified Majority

Designing a voting system for the Council one needs to set the threshold for the qualified majority. In general, this quota can be treated as a free parameter of the system and is often considered as a number to be negotiated. For political reasons one usually requires that the voting system should be *moderately conservative*, so one considers the quota in the wide range from 55 to 75 %.

However, designing the voting system based on the theory of Penrose, one can find a way to obtain a single number as the optimal value of the quota. In order to assure that the voting powers of all citizens in the "union" are equal one has to impose the requirement that the voting power of each member state should be proportional to the square root of the population of each state.

Let us analyze the problem of M members of the voting body, each representing a state with population N_i, $i = 1, \ldots, M$. Denote by w_i the voting weight attributed to each representative. We work with renormalized quantities, such that $\sum_{i=1}^{M} w_i = 1$. Assume that the decision of the voting body is taken, if the sum of the weights w_i of all members of the coalition exceeds the given quota q.

In the Penrose voting system one sets the voting weights proportional to the square root of the population of each state, $w_i \sim \sqrt{N_i}$ for $i = 1, \ldots, M$. For any level of the quota q one may compute numerically the power indices β_i. To characterize the overall representativeness of the voting system one may use various indices designed to quantify the resulting inequality in the distribution of power among citizens (Laruelle and Valenciano 2002). Analyzing the influence of the quota q for the average inequality of the voting power we are going to use the mean discrepancy Δ, defined as:

$$\Delta := \sqrt{\frac{1}{M}\sum_{i=1}^{M}(\beta_i - w_i)^2}\ , \tag{8}$$

Table 2 Optimal quota q_n for the Council of the European Union of M member states compared with predictions q_{av} of the approximate formula (16) and the lower bound q_{min} given in (10)

M	25	27	28	29	...	$M \to \infty$
$q_n(\%)$	62.16	61.58	61.38	61.32	...	50.0
$q_{av}(\%)$	61.28	60.86	60.66	60.48	...	50.0
$q_{min}(\%)$	60.00	59.62	59.45	59.28	...	50.0

The calculations of the optimal quotas for the EU were based upon the Eurostat data on the distribution of population for the EU-25 (2004) and the EU-27 (2010). The extended variant EU-28 contains EU-27 and Croatia, while EU-29 includes also Iceland

If the discrepancy Δ is equal to zero, the voting power of each state is proportional to the square root of its population. Under the assumption that the Penrose law is fulfilled, in such a case the voting power of any citizen in each state is the same.

In practice, the coefficient Δ will not be exactly equal to zero, but one may try to minimize this quantity. The optimal quota q_* can be defined as the quota for which the discrepancy Δ is minimal. Let us note, however, that this definition works fine for the Banzhaf index, while the dependence of the Shapley–Shubik index (Shapley and Shubik 1954) on the quota does not exhibit such a minimum.

Studying the problem for a concrete distribution of the population in the European Union, it was found (Słomczyński and Życzkowski 2004) that in these cases all M ratios β_i/w_i for $i = 1, \dots, M$, plotted as a function of the quota q, cross approximately near a single point. In other words, the discrepancy Δ at this critical point q_* is negligible. Numerical analysis allows one to conclude that this optimal quota is approximately equal to 62.0 % for the EU-25 (Słomczyński and Życzkowski 2004). At this very level of the quota the voting system can be considered as optimal, since the voting power of all citizens becomes equal. Performing detailed calculations one needs to care to approximate the square root function with a sufficient accuracy, since the rounding effects may play a significant role (Kurth 2007).

It is worth to emphasize that in general the value of the optimal quota decreases with the number of member states. For instance, in the case of the EU-27 is equal to 61.5 % (Życzkowski et al. 2006; Słomczyński and Życzkowski 2007), see Table 2. The optimal quota was also found for other voting bodies including various scenarios for an EU enlargement—see Leech and Aziz (2010). Note that the above results belong to the range of values of the quota for qualified majority, which are used in practice or recommended by experts.

4.1 Large Number of Member States and a Statistical Approximation

Further investigation has confirmed that the existence of such a critical point is not restricted to the concrete distribution of the population in European Union. On the contrary, it was reported for a model union containing M states with a random distribution of population (Słomczyński and Życzkowski 2004, 2006; Chang et al. 2006). However, it seems unlikely that we can obtain an analytical expression for the optimal quota in such a general case. If the number of member states is large enough one may assume that the distribution of the sum of the weights is approximately Gaussian (Owen 1975; Feix et al. 2007; Słomczyński and Życzkowski 2007). Such an assumption allowed us to derive an explicit approximate formula for the optimal quota for the Penrose square root voting system (Słomczyński and Życzkowski 2007)

$$q_{\mathrm{n}} := \frac{1}{2}\left(1 + \frac{\sqrt{\sum_{i=1}^{M} N_i}}{\sum_{i=1}^{M} \sqrt{N_i}}\right), \tag{9}$$

where N_i denotes the population of the i-th state. In practice it occurs that already for $M = 25$ this approximation works fine and in the case of the EU-25 gives the optimal quota with an accuracy much better than 1 %. Although the value of the optimal quota changes with M, the efficiency of the system, measured by the probability of forming the winning coalition, does not decrease if the union is enlarged. It was shown in Słomczyński and Życzkowski (2007) that, according to the central limit theorem, the efficiency of this system tends to approximately 15.9 % if $M \to \infty$.

It is not difficult to prove that for any fixed M the above expression attains its minimum if the population of each member state is the same, $N_i = \mathrm{const}\,(i)$. In this way one obtains a lower bound for the optimal quota as a function of the number of states (Słomczyński and Życzkowski 2007):

$$q_{\min} := \frac{1}{2}\left(1 + \frac{1}{\sqrt{M}}\right). \tag{10}$$

Note that the above bound decreases with the number of the states forming the union as $1/\sqrt{M}$ to 50 %. Such a behavior, reported in numerical analysis of the problem (Słomczyński and Życzkowski 2004, 2006; Chang et al. 2006) is consistent with the so-called Penrose limit theorem—see Lindner and Machover (2004).

4.2 *Optimal Quota Averaged over an Ensemble of Random States*

Concrete values of the optimal quota obtained by finding numerically the minimum of the discrepancy (8) for the EU-25 and the EU-27 (Słomczyński and Życzkowski 2004, 2006, 2010) are consistent, with an accuracy up to 2 %, with the data obtained numerically by averaging over a sample of random distribution of the populations of a fictitious union. This observation suggests that one can derive analytically an approximate formula for the optimal quota by averaging the explicit expression (9) over an ensemble of random populations N_i.

To perform such a task let us denote by x_i the relative population of a given state, $x_i = N_i / \sum_{i=1}^{M} N_i$. Since $\sqrt{N_i}/\sqrt{\sum_{i=1}^{M} N_i} = \sqrt{x_i}$ one can rewrite expression (9) in the new variables to obtain

$$q_{\mathrm{n}}(\vec{x}) = \frac{1}{2}\left(1 + \frac{1}{\sum_{i=1}^{M} \sqrt{N_i}/\sqrt{\sum_{i=1}^{M} N_i}}\right) = \frac{1}{2}\left(1 + \frac{1}{\sum_{i=1}^{M} \sqrt{x_i}}\right). \tag{11}$$

By construction, $\vec{x} = (x_1, \ldots, x_M)$ forms a probability vector with $x_i \geq 0$ and $\sum_{i=1}^{M} x_i = 1$. Hence the entire distribution of the population of the union is characterized by the M-point probability vector $\vec{x}$, which lives in an $(M-1)$ dimensional simplex Δ_M. Without any additional knowledge about this vector we can assume that it is distributed uniformly on the simplex,

$$P_D(x_1, \ldots, x_M) = \frac{1}{(M-1)!}\, \delta\left(1 - \sum_{i=1}^{M} x_i\right). \tag{12}$$

Technically it is a particular case of the *Dirichlet distribution*, written $P_D(\vec{x})$, with the Dirichlet parameter set to unity.

In order to get a concrete result one should then average expression (11) with the flat probability distribution (12). Result of such a calculation can be roughly approximated by substituting M-fold mean value over the Dirichlet measure, $M\langle\sqrt{x}\rangle_D$, instead of the sum into the denominator of the correction term in (11),

$$q_{\mathrm{av}}(M) := \langle q_{\mathrm{n}}\rangle_D \;\approx\; \frac{1}{2}\left(1 + \frac{1}{M\langle\sqrt{x}\rangle_D}\right). \tag{13}$$

The mean square root of a component of the vector $\vec{x}$ is given by an integral with respect to the Dirichlet distribution

$$\langle\sqrt{x}\rangle_D = \int_{\Delta_M} \sqrt{x_1}\, P_D(x_1, \ldots, x_M)\, dx_1 \cdots dx_M. \tag{14}$$

Instead of evaluating this integral directly, we shall rely on some simple fact from the physical literature. It is well known that the distribution of the squared absolute values of an expansion of a random state in an M-dimensional complex Hilbert space is given just by the flat Dirichlet distribution (see e.g. Bengtsson and Życzkowski 2006). In general, all moments of such a distribution where computed by Jones (1991). The average square root is obtained by taking his expression (26) and setting $d = M$, $l = 1$, $\nu = 2$ and $\beta = 1/2$. This gives the required average

$$\langle\sqrt{x}\rangle_D = \frac{\Gamma(M)\,\Gamma(3/2)}{\Gamma(M+1/2)} \sim \frac{\sqrt{\pi}}{2\sqrt{M}}. \tag{15}$$

Here Γ denotes the Euler gamma function and the last step follows from its Stirling approximation. Substituting the average $\langle\sqrt{x}\rangle_D$ into (13) we arrive at a compact expression

$$q_{\text{av}}(M) \approx \frac{1}{2} + \frac{1}{\sqrt{\pi M}} = \frac{1}{2}\left(1 + \frac{2}{\sqrt{\pi}}\frac{1}{\sqrt{M}}\right). \tag{16}$$

This approximate formula for the mean optimal quota for the Penrose voting system in a union of M random states constitutes the central result of this work. Note that this expression is averaged over all possible distributions of populations in the union, so it depends only on the size M of the union and on the form of averaging. The formula has a similar structure as the lower bound (10), but the correction term is enhanced by the factor $2/\sqrt{\pi} \approx 1.128$. In some analogy to the famous *Buffon's needle (or noodle) problem* (Ramaley 1969), the final result contains the number π—it appears in (16) as a consequence of using the normal approximation. The key advantage of the result (16) is due to its simplicity. Therefore, it can be useful in a practical case, if the size M of the voting body is fixed, but the weights of the voters (e.g. the populations in the EU) vary.

5 Concluding Remarks

In this work we review various arguments leading to the weighted voting system based upon the square root law of Penrose. However, the key result consists in an approximate formula for the mean optimal threshold of the qualified majority. It depends only on the number M of the states in the union, since the actual distribution of the population is averaged out.

Making use of this result we are in a position to propose a simplified voting system. The system consists of a single criterion only and is determined by the following two rules:

1. Each member of the voting body of size M is attributed his voting weight proportional to the square root of the population he represents.
2. The decision of the voting body is taken if the sum of the weights of members of a coalition exceeds the critical quota $q = 1/2 + 1/\sqrt{\pi M}$.

This voting system is based on a single criterion. Furthermore, the quota depends on the number of players only, but not on the particular distribution of weights of the individual players. This feature can be considered as an advantage in a realistic case, if the distribution of the population changes in time. The system proposed is objective and it cannot a priori handicap a given member of the voting body. The quota for qualified majority is considerably larger than 50 % for any size of the voting body of a practical interest. Thus the voting system is moderately conservative, as it should be. If the distribution of the population is known and one may assume that it is invariant in time, one may use a modified rule (2') and set the optimal quota according to the more precise formula (9).

Furthermore, the system is transparent: the voting power of each member of the voting body is up to a high accuracy proportional to his voting weight. However, as a crucial advantage of the proposed voting system we would like to emphasize its extendibility: if the size M of the voting body changes, all one needs to do is to set the voting weights according to the square root law and adjust the quota q according to the rule (2). Moreover, for a fixed number of players, the system does not depend on the particular distribution of weights. This feature is specially relevant for voting bodies in corporate management for which the voting weights may vary frequently.

It is our pleasure to thank W. Kirsch, M. Machover, F. Pukelsheim for fruitful discussions and to E. Ratzer for helpful correspondence. We are obliged to the anonymous referee for numerous comments and suggestions which allowed us to improve the article. Financial support by the European grant COCOS is gratefully acknowledged.

References

Ade, F. (2006). *Decision making in Europe: Were Spain and Poland right to stop the constitution in December 2003?* Preprint. http://congress.utu.fi/epcs2006/docs/D3_ade.pdf

Algaba, E., Bilbao, J. R., & Fernandez, J. R. (2007). The distribution of power in the European Constitution. *European Journal of Operational Research*, *176*, 1752–1766.

Andjiga, N.-G., Chantreuil, F., & Leppeley, D. (2003). La mesure du pouvoir de vote. *Mathematical Social Sciences*, *163*, 111–145.

Baldwin, R. E., & Widgrén, M. (2004). *Winners and losers under various dual majority rules for the EU Council of Ministers* (CEPR Discussion Paper No. 4450, Centre for European Policy Studies, Brussels 2004). http://www.cepr.org/pubs/dps/DP4450.asp

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

Bârsan-Pipu, N., & Tache, I. (2009). An analysis of EU voting procedures in the enlargement context. *International Advances in Economic Research*, *15*, 393–408.

Beisbart, C., & Bovens, L., & Hartmann, S. (2005). A utilitarian assessment of alternative decision rules in the Council of Ministers. *European Union Politics*, *6*, 395–419.

Bengtsson, I., & Życzkowski, K. (2006). *Geometry of quantum states.* Cambridge: Cambridge University Press.

Chang, P.-L., Chua, V. C. H., & Machover, M. (2006). LS Penrose's limit theorem: Tests by simulation. *Mathematical Social Sciences*, *51*, 90–106.

Feix, M. R., Lepelley, D., Merlin, V., & Rouet, J. L. (2007). On the voting power of an alliance and the subsequent power of its members. *Social Choice and Welfare*, *28*, 181–207.

Felsenthal, D. S., & Machover M. (1997). The weighted voting rule in the EUs Council of Ministers, 1958–1995: Intentions and outcomes. *Electoral Studies*, *16*, 33–47.

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

Felsenthal, D. S., & Machover, M. (1999). Minimizing the mean majority deficit: The second square-root rule. *Mathematical Social Science*, *37*, 25–37.

Felsenthal, D. S., & Machover, M. (2001). Treaty of Nice and qualified majority voting. *Social Choice and Welfare*, *18*, 431–464.
Gelman, A., Katz, J. M., & Tuerlinckx, F. (2002). The mathematics and statitics of voting power. *Statistical Science*, *17*, 420–435.
Gelman, A., Katz, J. M., & Bafumi, J. (2004). Standard voting power indexes do not work: An empirical analysis. *British Journal of Political Science*, *34*, 657–674.
Hosli, M. O. (2008). Council decision rules and European Union constitutional design. *AUCO Czech Economic Review*, *2* , 76–96.
Jones, K. R. W. (1991). Riemann-Liouville fractional integration and reduced distributions on hyperspheres. *Journal of Physics*, *A 24*, 1237–1244.
Kirsch, W. (2007). On Penrose's square-root law and beyond. *Homo Oeconomicus*, *24*, 357–380.
Kirsch, W. (2010). The distribution of power in the Council of Ministers of the European Union. In M. Cichocki, & K. Życzkowski (Eds.), *Institutional design and voting power in the European Union* (pp. 93–107). Farnham: Ashgate Publishing Group.
Kirsch, W., Słomczyński, W., & Życzkowski, K. (2007). Getting the votes right. *European Voice*, 3–9, 12.
Kurth, M. (2007). *Square root voting in the Council of the European Union: Rounding effects and the jagiellonian compromise*. Preprint [math.GM 0712.2699].
Laruelle, A., & Valenciano, F. (2002). Inequality among EU citizens in the EU's Council decision procedure. *European Journal of Political Economy*, *18*, 475–498.
Laruelle, A., & Valenciano, F. (2008). *Voting and collective decision-making. Bargaining and power*. Cambridge: Cambridge University Press.
Laruelle, A., & Widgrén, M. (1998). Is the allocation of voting power among the EU states fair? *Public Choice*, *94*, 317–339.
Leech, D. (2002). Designing the voting system for the Council of the EU. *Public Choice*, *113*, 437–464.
Leech, D., & Aziz, H. (2010). The double majority voting rule of the EU reform treaty as a democratic ideal for an enlarging Union: An appraisal using voting power analysis. In M. Cichocki, & K. Życzkowski (Eds.), *Institutional design and voting Power in the European Union* (pp. 59–73). Farnham: Ashgate Publishing Group.
Leech, D., & Machover, M. (2003). Qualified majority voting: The effect of the quota. In M. Holler, et al. (Eds.), *European Governance, Jahrbuch für Neue Politische Ökonomie* (pp. 127–143). Tübingen: Mohr Siebeck.
Lindner, I., & Machover, M. (2004). LS Penrose's limit theorem: Proof of some special cases. *Mathematical Social Sciences*, *47*, 37–49.
Machover, M. (2010). Penrose's square root rule and the EU Council of the Ministers: Significance of the quota. In M. Cichocki, & K. Życzkowski (Eds.), *Institutional design and voting power in the European Union* (pp. 35–42). Farnham: Ashgate Publishing Group.
Moberg, A. (2010). Is the double majority really double? The voting rules in the Lisbon Treaty. In M. Cichocki, & K. Życzkowski (Eds.), *Institutional design and voting power in the European Union* (pp. 19–34). Farnham: Ashgate Publishing Group.
Morriss, P. (1987). *Power: A philosophical analysis*. Manchester: Manchester University Press; 2nd ed. 2002.
Owen, G. (1975). Multilinear extensions and the Banzhaf value. *Naval Research Logistics Quaterly*, *22*, 741–750.
Pajala, A., & Widgrén, M. (2004). A priori versus empirical voting power in the EU Council of Ministers. *European Union Politics*, *5*, 73–97.
Penrose, L. S. (1946). The elementary statistics of majority voting. *Journal of the Royal Statistical Society*, *109*, 53–57.
Penrose, L. S. (1952). *On the objective study of crowd behaviour*. London: H.K. Lewis & Co.
Pöppe, Ch. (2007, August). Die Quadratwurzel, das Irrationale und der Tod. *Spektrum der Wissenschaft*, 102–105.
Pukelsheim, F. (2007, Juni 20). Der Jagiellonische Kompromiss. *Neue Züricher Zeitung*.

Pukelsheim, F. (2010). Putting citizens first: Representation and power in the European Union. In M. Cichocki, & K. Życzkowski (Eds.), *Institutional design and voting power in the European Union* (pp. 235–253). Farnham: Ashgate Publishing Group.
Ramaley, J. F. (1969). Buffon's noodle problem. *American Mathematical Monthly*, *76*, 916–918.
Shapley, L. S., & Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American Science Review*, *48*, 787–792.
Słomczyński, W., & Życzkowski, K. (2004, May). *Voting in the European Union: The square root system of Penrose and a critical point*. Preprint cond-mat.0405396.
Słomczyński, W., & Życzkowski, K. (2006). Penrose voting system and optimal quota. *Acta Physica Polonica*, *B 37*, 3133–3143.
Słomczyński, W., & Życzkowski, K. (2007). From a toy model to the double square root voting system. *Homo Oeconomicus*, *24*, 381–399.
Słomczyński, W., & Życzkowski, K. (2010). Jagiellonian compromise - an alternative voting system for the Council of the European Union. In M. Cichocki, & K. Życzkowski (Eds.), *Institutional design and voting power in the European Union* (pp. 43–57). Farnham: Ashgate Publishing Group.
Życzkowski, K., Słomczyński, W., & Zastawniak, T. (2006). Physics for fairer voting. *Physics World*, *19*, 35–37.

The Fate of the Square Root Law for Correlated Voting

Werner Kirsch and Jessica Langner

1 Introduction

In this paper, we consider two-tier voting systems. The first level of such a systems usually consists of the voters in a country or an association of countries. The voters in each constituency (or member country) are represented by a delegate in the second level voting system, the council. Delegates in the council are given a voting weight which as a rule depends on the population of the constituency they represent.

Examples of such two-tier voting systems are the Council of Ministers of the European Union, the Electoral College in the USA and the "Bundesrat", the state chamber of Germany's parliamentary system. In each case we assume that the representatives vote according to the majority vote in their respective constituency.

What is a fair voting weight for a delegate in a council? This question arises immediately in all these examples. It seems self-evident that for a fair voting system the voting outcome in the council should agree with the result of a popular vote. The US presidential elections 2000 show that this is not always the case. While Al Gore won the public vote the majority in the Electoral College elected George W. Bush as the 43rd president of the USA. The difference between the voting result in the council and the public vote is called the "democracy deficit".

In fact, it is not hard to see, that *no* voting system for the council can *guarantee* that the vote in the council and the public vote agree. In other words, no matter how we choose the voting weights for the council members, the democracy deficit cannot be zero for *all* possible distributions of "yes"- and "no"-votes among the voters. Thus, the best one can do is to minimize the expected democracy deficit, i.e., the difference between the vote in the council and popular vote. Obviously, the term "expected" needs a careful interpretation. If one assumes that all voters

---

W. Kirsch (✉) • J. Langner
Fakultät für Mathematik und Informatik, FernUniversität in Hagen, 58084 Hagen, Germany
e-mail: werner.kirsch@fernuni-hagen.de; jessica.langner@fernuni-hagen.de

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare, DOI 10.1007/978-3-319-05158-1_9,

cast their votes independently of each other then one can show that the expected democracy deficit is minimized if the voting weight of a representative is chosen proportional to the *square root* $\sqrt{N_\nu}$ of the population (N_ν) of the respective country (with number ν).

This is (one version of) the celebrated "square root law" by Penrose (see Felsenthal and Machover 1998; Penrose 1946). In this paper, we go beyond the square root law by dropping the assumption of the voters' independence. We apply two different schemes to model the correlation between the voters. In our main model we assume that the voters are influenced by a "common belief" of the society or—which is the same, technically speaking—by a strong group of opinion makers. We call this system the CBM (for "common belief model" or "collective bias model") (see Kirsch 2007). The CBM can be looked upon as a generalization of a model proposed by Straffin (1977) in connection with the Shapley–Shubik power index (see Shapley and Shubik 1954). The other model we look at takes into account that voters influence each other. It is based on a model (the Curie–Weiss Model) for ferromagnetic behaviour taken from statistical physics (see Kirsch 2007 and cf. Ellis 1985; Thompson 1972).

If we assume that the voters in different countries vote independently of each other, we can compute the optimal voting weights in terms of the expected margins of the voting outcome in the countries. For the CBM the optimal weights are proportional to the population N_ν. We also compute the expected democracy deficit for these models (for large N_ν).

Under the assumption that the voters influence each other also across country borders (according to the CBM) we can also compute the expected democracy deficit asymptotically. It turns out that in this case any voting weight is as good as any other one. In other words, on an asymptotical scale *any* distribution of voting weights is close to optimal.

2 The General Model

We consider a situation where M states (countries, constituencies) form a federation. The states are labeled by Greek characters, e.g., $\nu, \kappa, \ldots$. The number of voters (population) of the state ν is denoted by N_ν. Consequently, the total population of the union is given by $N = \sum_{\nu=1}^{M} N_\nu$.

We denote the vote of the voter i in state ν by $X_{\nu i}$. This voter may vote either "yes", in which case we set $X_{\nu i} = 1$ or "no" encoded as $X_{\nu i} = -1$. Consequently, the result of a simple majority voting in the state ν is represented by the sum $S_\nu = \sum_{i=1}^{N_\nu} X_{\nu i}$. A voting in that state is affirmative if $S_\nu > 0$. For the simplicity of notation and to avoid nonsignificant technicalities we assume that all N_ν are odd numbers, this excludes a draw described by $S_\nu = 0$.

We denote the voting decision in the state ν by $\chi_\nu = \chi_\nu(S_\nu)$ which we set equal to 1 if $S_\nu > 0$ and equal to -1 if $S_\nu \leq 0$. Thus, the representative of state ν will vote "yes" if $\chi_\nu = 1$ and "no" if $\chi_\nu = -1$. For later use we note that $\chi_\nu S_\nu = |S_\nu|$.

If we denote the voting weight for state ν in the council by g_ν then the voting result in the council is given by

$$C = \sum_{\nu=1}^{M} g_\nu \, \chi_\nu \,. \tag{1}$$

This voting result has to be compared with the popular vote given by

$$P = \sum_{\nu=1}^{M} S_\nu \,. \tag{2}$$

We call the absolute value of the difference between C and P the *democracy deficit* and denote it by Δ

$$\Delta \;=\; |\, C - P \,| \tag{3}$$

$$= \Big| \sum_{\nu=1}^{M} g_\nu \, \chi_\nu \;-\; \sum_{\nu=1}^{M} S_\nu \Big| . \tag{4}$$

The democracy deficit Δ depends explicitly on the voting weights $g_1, \ldots, g_M$. The voting weights should be chosen in such a way that the democracy deficit is as small as possible.

The voting results $X_{\nu i}$ are the voter's reaction on a particular proposal ω. Hence, the democracy deficit Δ depends on the given proposal ω as well. It is easy to choose the weights g_ν such that Δ vanishes for a *given* proposal. But our goal is to optimize the weights in such a way that Δ is small for *most* proposals. Thus, we look at the *expected* value of Δ^2, denoted by

$$\mathbb{D} \;:=\; \mathbb{E}\Big(\Delta^2\Big) . \tag{5}$$

We will call $\mathbb{D}$ the *expected democracy deficit* in the following (instead of the correct but clumsy "expected square of the democracy deficit").

By looking at expectation values we regard the proposals as random input to the voting system. Hence the probability that the next proposal to the system is a particular proposal ω is determined by a probability rule. We assume that there is no bias to certain proposals, in particular any proposal and its counterproposal have the same probability.

The voting system reacts in a deterministic (and rational) way to this random input. The voting results as well as the democracy deficit are therefore (otherwise deterministic) functions of the random input, the proposal. The voting outcome is a vector in the space $\Omega = \{-1, 1\}^N$, where N is the total number of voters and the probability distribution of the proposals equips Ω with probability distribution $\mathbb{P}$ as well, namely the probability of a given outcome $(X_1, \ldots, X_N)$ is the probability of all proposals ω that lead to that outcome. Since the voters react rationally they vote -1 on the opposite to a proposal they would favour and vice versa. Hence the probability distribution $\mathbb{P}$ satisfies

$$\mathbb{P}(X_1, \ldots, X_N) = \mathbb{P}(-X_1, \ldots, -X_N)\,. \tag{6}$$

We call such a measure a *voting measure*. For any voting measure we have $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}$, but probabilities concerning more than one voter, like $\mathbb{P}(X_1 = 1 \text{ and } X_2 = 1)$ cannot be computed from the mere assumption that $\mathbb{P}$ is a voting measure. Such events concern the correlation structure of the measure and they have yet to be fixed depending on the situation at hand. One possible specification is the assumption that all voters act independently of each other. This leads to the property that

$$\mathbb{P}(X_1 = 1 \text{ and } X_2 = 1) \;=\; \mathbb{P}(X_1 = 1) \cdot \mathbb{P}(X_2 = 1) \;=\; \tfrac{1}{4}\,.$$

More generally, under the assumption of independence we have

$$\mathbb{P}\big(X_1 = \xi_1, X_2 = \xi_2, \ldots, X_N = \xi_N\big) = \frac{1}{2^N} \tag{7}$$

for any $\xi_1, \ldots, \xi_N \in \{-1, 1\}$. The voting measure describes the mutual influence of the voters on each other, mathematically speaking it describes the correlation structure of the voting system. The above example describes *independent* voters—in some sense the classical case of the theory. An extreme case is given by the measure $\mathbb{P}_u$

$$\begin{aligned}\mathbb{P}_u\big(X_1 = 1, X_2 = 1, \ldots, X_N = 1\big) &= \mathbb{P}_u\big(X_1 = -1, X_2 = -1, \ldots, X_N = -1\big)\\ &= \frac{1}{2}\,.\end{aligned} \tag{8}$$

For this (rather boring) voting measure the only possible outcomes are the unanimous votes, it represents total (positive) correlation.

If $\mathbb{P}$ is a voting measure, we denote the expectation value with respect to $\mathbb{P}$ by $\mathbb{E}$, as was already anticipated in (5). Since we assume that the numbers N_ν are odd, it follows that $S_\nu \neq 0$. From this we conclude that $\mathbb{E}(\chi_\nu) = 0$ for any voting measure.

3 Optimal Weights for Independent States

We begin by determining optimal weights, under the assumption that voters in different states are independent. Thus, we assume that the random variables $X_{\nu i}$ and $X_{\kappa j}$ are independent for $\nu \neq \kappa$.

We want to minimize the function

$$\begin{aligned}\mathbb{D}(\gamma_1,\ldots,\gamma_M) &= \mathbb{E}\big(\,\Delta(\gamma_1,\ldots,\gamma_M)^2\big)\\ &= \sum_{\nu,\kappa=1}^{M}\Big(\gamma_\nu\gamma_\kappa\mathbb{E}\big(\chi_\nu\chi_\kappa\big)-2\gamma_\nu\mathbb{E}\big(\chi_\nu S_\kappa\big)+\mathbb{E}\big(S_\nu S_\kappa\big)\Big)\,. \qquad (9)\end{aligned}$$

The function $\mathbb{D}(\gamma_1,\ldots,\gamma_M)$ is a measure for the expected democracy deficit for voting weights $\gamma_1,\ldots,\gamma_M$.

By the assumption of independent *states* we can conclude that

$$\mathbb{E}\big(\chi_\nu\chi_\kappa\big) = \mathbb{E}\big(\chi_\nu\big)\,\mathbb{E}\big(\chi_\kappa\big)=0 \qquad \text{for } \nu\neq\kappa\,, \qquad (10)$$

$$\mathbb{E}\big(\chi_\nu S_\kappa\big) = \mathbb{E}\big(\chi_\nu\big)\,\mathbb{E}\big(S_\kappa\big)=0 \qquad \text{for } \nu\neq\kappa\,, \qquad (11)$$

and

$$\mathbb{E}\big(S_\nu S_\kappa\big) = \mathbb{E}\big(S_\nu\big)\,\mathbb{E}\big(S_\kappa\big)=0 \qquad \text{for } \nu\neq\kappa\,. \qquad (12)$$

Moreover, we have $\chi_\nu^2 = 1$ and $\chi_\nu S_\nu = |S_\nu|$, thus

$$\mathbb{D}(\gamma_1,\ldots,\gamma_M) = \sum_{\nu=1}^{M}\Big(\gamma_\nu^2 \;-\; 2\gamma_\nu\mathbb{E}\big(|S_\nu|\big) \;+\; \mathbb{E}\big(S_\nu^2\big)\Big)\,. \qquad (13)$$

It is not hard to find the minimizing weights g_ν (by the usual procedure: find the zeros of the derivative), in fact: the weights $g_1,\ldots,g_M$ which minimize the function $\mathbb{D}$ are given by

$$g_\nu = \mathbb{E}\big(\big|S_\nu\big|\big)\,. \qquad (14)$$

This result has a very intuitive interpretation. The quantity S_ν is the difference between the "yes"-votes and the "no"-votes, so $|S_\nu|$ describes the margin of the voting outcome, i.e., the surplus of votes of the winning party. Therefore, the optimal weights g_ν for the state ν are given by the expected margin of a vote in that state. In fact, the delegate of state ν does not represent the opinion of *all* voters in this state, but only those who agree with the majority, he or she acts against the will of the minority, so as a net result the delegate just represents the margin.

We can also compute the expected democracy deficit $\mathbb{D}$ for the optimal weights $g_1,\ldots,g_M$

$$\mathbb{D}(g_1,\ldots,g_M) = \sum_{\nu=1}^{M}\Big(\mathbb{E}\big(\big|S_\nu\big|^2\big)-\mathbb{E}\big(\big|S_\nu\big|\big)^2\Big) = \sum_{\nu=1}^{M}\mathbb{V}\big(\big|S_\nu\big|\big) \qquad (15)$$

where $\mathbb{V}(|S_\nu|)$ denotes the variance of the random quantity $|S_\nu|$.

We emphasize that we did not yet make assumptions about the correlation structure of voters *inside* a country. Of course, the numerical evaluation of the optimal weights and minimal democracy deficit requires further assumptions on the correlation between voters.

4 Independent Voters

In this section we assume that all voters act independently of each other, in mathematical terms: all random variables $X_{\nu i}$ are independent of each other. Under this assumption we can compute the optimal weight $g_\nu = \mathbb{E}(|S_\nu|)$ as well as the minimal expected democracy deficit.

For the independent random variables $X_{\nu i}$ we have the central limit theorem, namely the weighted sums

$$\frac{1}{\sqrt{N_\nu}} S_\nu := \frac{1}{\sqrt{N_\nu}} \sum_{i=1}^{N_\nu} X_{\nu i} \tag{16}$$

are asymptotically distributed for large N_ν according to a standard normal distribution (cf. Lamperti 1996). From this it follows that for large N_ν

$$\mathbb{E}\big(|S_\nu|\big) \approx \frac{\sqrt{2}}{\sqrt{\pi}} \sqrt{N_\nu}\,, \tag{17}$$

$$\mathbb{E}\big(|S_\nu|^2\big) \approx \sqrt{N_\nu}\,, \tag{18}$$

and

$$\mathbb{V}\big(|S_\nu|\big) \approx \frac{\pi - 2}{\pi} N_\nu\,. \tag{19}$$

We conclude that the optimal weight for independent voters is proportional to the square root of the population. This is exactly the content of the square root law by Penrose (see Penrose 1946; Felsenthal and Machover 1998).

The above formulae also allow us to evaluate the minimum of the expected democracy deficit

$$\mathbb{D}(g_1, \ldots, g_M) \approx \frac{\pi - 2}{\pi} N\,. \tag{20}$$

This implies that the *expected democracy deficit per voter*, namely

$$\mathbb{E}\left(\left(\frac{\Delta}{N}\right)^2\right) \tag{21}$$

converges to zero as N becomes large (with convergence rate $\frac{1}{N}$).

5 The Collective Bias Model

Now, we introduce and discuss a model for collective behaviour of voters. The basic idea is that there is a mainstream opinion, e.g., a common belief due to the country's tradition or the influence of opinion makers. For a given proposal ω we model this "common belief" by a value $\zeta \in [-1, 1]$ which depends on the proposal at hand. The value $\zeta = 1$ means there is such a strong common belief in favor of the proposal that all voters will vote "yes", $\zeta = -1$ means all voters will vote "no". In general, ζ denotes the expected outcome of the voting, i.e., $\mathbb{E}(X_{\nu i})$. The voting results $X_{\nu i}$ themselves fluctuate around this value randomly.

Let us be more precise about this. Suppose the voting results are $X_1, \ldots, X_N$ (where we dropped the index ν for notational simplicity). Let μ be a measure on $[-1, 1]$, which is the distribution of the common belief value ζ, that is $\mu(]a, b[)$ is the probability that the value ζ is between a and b. Let P_ζ be the probability measure on $\{-1, 1\}$ with

$$P_\zeta(X_1 = 1) = p_\zeta = \tfrac{1}{2}(1 + \zeta)\,,$$

so that

$$E_\zeta(X_1) := P_\zeta(X_1 = 1) - P_\zeta(X_1 = -1) = p_\zeta - (1 - p_\zeta) = \zeta\,.$$

For a given value of ζ we set

$$\mathcal{P}_\zeta(\xi_1, \ldots, \xi_N) = \prod_{i=1}^{N} P_\zeta(\xi_i)\,. \tag{22}$$

For any $\zeta \in [-1, 1]$ the expression $\mathcal{P}_\zeta$ is a probability distribution on $\Omega = \{-1, 1\}^N$. We define the collective bias measure $\mathbb{P}_\mu$ with respect to μ as

$$\mathbb{P}_\mu(X_1 = \xi_1, \ldots, X_N = \xi_N) \;:=\; \int \mathcal{P}_\zeta(\xi_1, \ldots, \xi_n)\, d\mu(\zeta)\,. \tag{23}$$

Note, that $\mathcal{P}_\zeta$ is *not* a voting measure (unless $\zeta = \frac{1}{2}$). However $\mathbb{P}_\mu$ *is* a voting measure if μ is invariant under sign change, i.e., $\mu(]a, b[) = \mu(]-b, -a[)$. We call μ the *bias measure*.

If the measure μ is concentrated in 0, then $\mathbb{P}_\mu$ makes the voting results X_i independent, thus we are in the case of Sect. 4. If μ is the uniform distribution on $[-1, 1]$ (that is every point is equally likely), then the corresponding measure was already considered by Straffin (1977) where he established an intimate connection of this model to the Shapley–Shubik index. In a similar way, the Penrose–Banzhaf measure is connected with the model of independent voters.

The Collective Bias Model (CBM) can be looked upon as a model for spins in statistical mechanics. There the voters are replaced with elementary magnets (spins) which can be directed upwards ($X_i = 1$) or downwards ($X_i = -1$). In this language the Collective Bias Model describes spins which do not interact with each other but are influenced by an exterior magnetic field, namely the collective bias ζ.

In the papers (Kirsch 2007; Kirsch and Langner 2014; Langner 2012) we investigate also another model for collective voting behaviour which comes directly from statistical physics, the Curie–Weiss Model (CWM). In this model the spins (voters) influence each other by an interaction which makes spins to prefer to be directed parallel to the others. For voting this means that voters prefer to agree to the other voters. The Curie–Weiss Model is a very interesting tool to investigate collective behaviour. However, it is technically more involved than the other models we discuss. Therefore, we will mention it only rather briefly and refer to the papers mentioned above for more details.

Let us define

$$H(X_1, \dots, X_N) = -\frac{1}{N}\Big(\sum_{i=1}^{N} X_i\Big)^2 . \tag{24}$$

This is the *energy function* for the spin configuration $X_1, \dots, X_N$. We use this to define measures

$$Q_\beta(X_1, \dots, X_N) = e^{-\beta H(X_1,\dots,X_N)} \tag{25}$$

where $\beta \in]0, \infty[$ is the inverse temperature in statistical physics. As a rule, Q_β is not a probability measure, so we normalize it by dividing through its total mass Z and set

$$P_\beta(X_1, \dots, X_N) = \frac{e^{-\beta H(X_1,\dots,X_N)}}{Z} . \tag{26}$$

This is the Curie–Weiss measure for inverse temperature β. The parameter β measures the strength of the interaction between the voters. The extreme case $\beta = 0$ corresponds to the model of independent voters, the other extreme $\beta = \infty$ describes the case of the measure $\mathbb{P}_u$ defined in (8) for unanimous voting.

6 Optimal Weights for the Collective Bias Model

Let us now suppose that voters in different countries are independent, but voting inside the countries follows the CBM with bias measure μ. According to Sect. 3 in this case the optimal weights are given by

$$g_\nu = \mathbb{E}_\mu\big(|S_\nu|\big)\,. \tag{27}$$

For large N_ν we have

$$g_\nu = \mathbb{E}_\mu\big(|S_\nu|\big) \;=\mu_1\,N_\nu \tag{28}$$

where $\mu_1 = \int|\zeta|d\mu(\zeta)$ is the first absolute moment of μ. Note, that for any probability measure μ the quantity μ_1 is non zero, except for the case $\mu = \delta_0$, the measure is concentrated at the point 0. This means that the optimal weights for a council are proportional to the population of the respective country if the voters can be described by a CBM. This also includes the Straffin case (μ is the uniform distribution), which corresponds to the Shapley–Shubik power index.

The only exception from proportionality is the case $\mu = \delta_0$ corresponding to independent voting (the Penrose–Banzhaf case), where the square root law applies.

We mention that there is a "phase transition" for the Curie–Weiss Model if we vary β from 0 to ∞, namely

$$g_\nu = \mathbb{E}_\beta\big(|S_\nu|\big) = \begin{cases} \frac{\sqrt{2}}{\sqrt{\pi}\sqrt{1-\beta}}\,\sqrt{N_\nu}, & \text{for } \beta < 1; \\ C\,N_\nu^{\frac{3}{4}}, & \text{for } \beta = 1; \\ C(\beta)\,N_\nu, & \text{for } \beta > 1. \end{cases} \tag{29}$$

The constant $C(\beta)$ converges to 0 as $\beta \searrow 1$ and to 1 as $\beta \nearrow \infty$.

7 Democracy Deficit for the Collective Bias Model

Given the optimal weights (28) for the CBM (and independent states) we can compute (the asymptotic behaviour of) the expected democracy Deficit $\mathbb{D}_\mu$

$$\mathbb{D}_\mu = \sum_{\nu=1}^{M} \mathbb{V}\big(|S_\nu|\big) \;\approx\; (\mu_2 - \mu_1^2)\,N^2 \tag{30}$$

where $\mu_1 = \int |\zeta| d\mu(\zeta)$ and $\mu_2 = \int |\zeta|^2 d\mu(\zeta)$. Note that $\mu_2 - \mu_1^2 \neq 0$ unless μ is concentrated in at most two points. It follows that the expected democracy deficit per voter, i.e.,

$$\mathbb{E}_\mu\left(\left(\frac{\Delta}{N}\right)^2\right)$$

converges to a *positive* constant as the N_ν tend to infinity (in a uniform way, i.e., $N_\nu = \alpha_\nu N$).

It is interesting to remark that the expected democracy deficit per voter converges also to a constant if we choose a *non optimal* voting weight, like for instance $g_\nu \sim \sqrt{N_\nu}$ or $g_\nu = 1$ for all ν. This constant will in general be larger than the one for the optimal weights, but the order of magnitude of $\mathbb{D}$ is *not* changed.

For the Curie–Weiss Model the expected democracy deficit per voter converges to zero (for $\beta \neq 1$ even with rate $\frac{1}{N}$).

8 A Model with Global Collective Behaviour

So far we have always assumed that voter in different states act independently. In this section we consider the case of collective behaviour across country borders. We assume that *all* voters act according to the Collective Bias measure $\mathbb{P}_\mu$. This means there is a common belief, expressed through the measure μ, for all voters in the union.

Then, the formulae (10)–(13) are no longer valid. In fact, determining the optimal voting weights requires to solve a rather complicated system of M dependent linear equations. Instead of doing this we try to look at the democracy deficit directly. It turns out that for large N_ν we have for any ν, κ

$$\mathbb{E}_\mu\big(\chi_\nu \chi_\kappa\big) \approx 1\,, \tag{31}$$

$$\mathbb{E}_\mu\big(\chi_\nu S_\kappa\big) \approx \mathbb{E}_\mu\big(\big|S_\kappa\big|\big) \approx \mu_1\, N_\kappa \tag{32}$$

and

$$\mathbb{E}\big(S_\nu S_\kappa\big) \approx \mu_2\, N_\nu\, N_\kappa\,. \tag{33}$$

Inserting these terms into the expression for $\mathbb{D}$ we obtain

$$\begin{aligned}\mathbb{D}(g_1,\ldots,g_M) \;=\; &\sum_{\nu,\kappa=1}^{M} \mathbb{E}_\mu(\chi_\nu \chi_\kappa)\, g_\nu g_\kappa \;-\; 2\sum_{\nu=1}^{M} g_\nu \sum_{\kappa=1}^{M} \mathbb{E}_\mu(\chi_\nu S_\kappa)\\ &+ \sum_{\nu,\kappa=1}^{M} \mathbb{E}_\mu(S_\nu S_\kappa)\end{aligned}$$

$$
\begin{aligned}
&\approx \sum_{\nu,\kappa=1}^{M} g_\nu g_\kappa \;-\; 2\sum_{\nu=1}^{M} g_\nu \sum_{\kappa=1}^{M} \mu_1\, N_\kappa \;+\; \sum_{\nu,\kappa=1}^{M} \mu_2\, N_\nu N_\kappa \\
&= \Big(\sum_{\nu=1}^{M} g_\nu\Big)^2 \;-\; 2\mu_1 \big(\sum_{\nu=1}^{M} g_\nu\big)\, N \;+\; \mu_2\, N^2 \\
&= G^2 \;-\; 2\mu_1\, G \;+\; \mu_2\, N^2\,. \qquad (34)
\end{aligned}
$$

This last expression depends only on the sum $G = \sum_{\nu=1}^{M} g_\nu$ of the voting weights and *not* on the single weight g_ν. This means that for large N_ν the asymptotic value of $\mathbb{D}$ does not depend on the way the weights are distributed among the member states of the union. The minimal value of $\mathbb{D}$ is obtained by choosing $G = \mu_1 N$ independently of the values of the particular weight g_ν. We also note that the value of G has no real meaning, since we don't change the voting system at all if we multiply all weights (and the quota) by the same number $C > 0$.

Finally, we remark that the somewhat hand waving arguments in (34) need a careful mathematical interpretation. A precise formulation gives:

$$\lim_{N\to\infty} \mathbb{E}_\mu\Big(\Big(\frac{\Delta(g_1,\ldots,g_M)}{N}\Big)^2\Big) = \mu_2 - \mu_1^2 \qquad (35)$$

for $G = \sum_{\nu=1}^{M} g_\nu = \mu_1 N$ and

$$\liminf_{N\to\infty} \mathbb{E}_\mu\Big(\Big(\frac{\Delta(g_1,\ldots,g_M)}{N}\Big)^2\Big) \;\geq\; \mu_2 - \mu_1^2 \qquad (36)$$

for any arbitrary choice of g_ν. This result can be interpreted in the following way: If there is a strong common belief in the union across border lines then it doesn't matter how one distributes the voting weights in the council.

Acknowledgements This paper has been presented at the Leverhulme Trust sponsored Voting Power in Practice Symposium held at the London School of Economics, 20–22 March 2011.

References

Ellis, R. (1985). Entropy, large deviations and statistical mechanics. *Grundlehren der mathematischen Wissenschaften* (Vol. 271). Berlin: Springer.

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

Kirsch, W. (2007). On Penrose's squareroot law and beyond. *Homo Oeconomicus*, *24*(3, 4), 357–380.

Kirsch, W., & Langner, J. (2014). Mathematical theory of correlated voting (in preparation).

Lamperti, J. W. (1996). *Probability: A survey of the mathematical theory* (2nd ed.). New York: Wiley.

Langner, J. (2012). Fairness, Efficiency and Democracy Deficit: Combinatorial Methods and Probabilistic Analysis on the Design of Voting Systems. Südwestdeutscher Verlag für Hochschulschriften.

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

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

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

Thompson, C. (1972). *Mathematical statistical mechanics*. Princeton: Princeton University Press.

The Mean Voter, the Median Voter, and Welfare-Maximizing Voting Weights

Nicola Maaser and Stefan Napel

1 Introduction

An important application of voting power analysis (see Felsenthal and Machover 1998, for a comprehensive overview) concerns the question of how voting weights should be assigned in *two-tier voting systems*. At the bottom tier, countries, states, districts, or other kinds of constituencies each elect a representative who will on their behalf cast a block vote in a top tier assembly or council. The Council of Ministers of the European Union (EU) is one of the most prominent examples of such a system, and much research on fair or optimal design of voting rules has been stimulated by successive EU enlargements. Other examples include the International Monetary Fund (see Leech and Leech 2009), the German Bundesrat and, with inessential qualifications, the US Electoral College.

We are grateful for constructive discussion at the Leverhulme Trust's *Voting Power in Practice Symposium* 2011.

N. Maaser
Department of Economics & ZeS, University of Bremen, Mary-Somerville-Str. 5, 28359 Bremen, Germany
e-mail: maaser@uni-bremen.de

S. Napel (✉)
Department of Economics, University of Bayreuth, Universitätsstr. 30, 95447 Bayreuth, Germany

Public Choice Research Centre, Turku, Finland
e-mail: stefan.napel@uni-bayreuth.de

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare, DOI 10.1007/978-3-319-05158-1_10,

In order to evaluate the design of two-tier voting systems a multitude of normative criteria can be brought to bear. Depending on the application at hand, desirable features include the responsiveness of collective decisions to individual preference changes, the capability to reach a decision, or equality of representation. From the perspective of mainstream economics, *utilitarian welfare* is a particularly prominent criterion (see Harsanyi 1955 and, e.g., Barberà and Jackson 2006). In particular, if the design of a two-tier voting system maximizes the total expected utility of the citizens, it is Pareto efficient: no other system can raise expected utility of some citizens without lowering it for others.

In this paper, we study the relationship between the allocation of block voting rights, i.e., the *voting weights* of constituency representatives, and the utilitarian welfare that is induced by the outcomes of a two-tier decision making process. We consider a model in which the feasible policy alternatives constitute a finite or infinite real interval. Voter preferences are assumed to be single-peaked, i.e., an individual's utility from a particular collective decision is strictly decreasing in distance to the respective voter's ideal point. These ideal points are conceived of as random variables with an identical continuous distribution for all citizens.

For ease of exposition, we suppose that each constituency comprises an odd number of voters. Then we assume, first, that the policy advocated by the single representative of any given constituency is congruent with the ideal point of the respective *constituency's median voter*. Second, the decision which is taken at the top tier is identified with the position of the *pivotal representative*. This representative is determined by the given allocation of voting weights and a 50 % decision quota together with the policy positions of all delegates. It corresponds to the *weighted median amongst the delegates* and to the *core* of the spatial voting game in the assembly. Consideration of the respective, generically single-valued core provides a short-cut to the equilibrium outcome of various conceivable negotiation protocols, which might structure strategic bargaining at the council level.[1] As long as the weighted median of delegates who represent their constituencies' median voters is a reasonable approximation for the outcomes generated by the two-tier voting system, the actual processes of preference aggregation within the council and within the constituencies can remain unspecified. In particular, the latter could differ across constituencies.

We take a set of differently sized constituencies as given and seek to find the weight allocation rule which maximizes total expected utility. We presume that the preferences over policy outcomes have the same cardinal intensity across voters and distinguish between two utility specifications. Namely, each voter's cardinal utility function decreases either linearly or quadratically in the distance between the individual's ideal point and the collective policy outcome. The former specification corresponds to voters who are risk neutral, i.e., who are indifferent between facing distances x and y to their ideal points with probabilities p and $1 - p \in (0, 1)$,

[1] See, e.g., Cho and Duggan (2009).

or suffering the expected distance $px + (1 - p)y$ for sure. The latter specification describes risk-averse individuals.

Standard results from statistics imply that the position of the *population's median voter* maximizes total expected utility for the linear specification, whereas the position of the *population's mean voter* maximizes it for the quadratic specification. It is, however, a non-trivial question how the best estimates for the sample median or sample mean, respectively, can be obtained by computing a weighted median of the medians of differently sized sub-samples. We are not aware of—and have unfortunately neither been able to obtain—general analytical results on this issue. We, therefore, conduct extensive computer simulations.

The main finding of our Monte Carlo analysis is that a *square root rule* should be used in order to allocate voting weights if all citizens are a priori identical in a strong sense: namely, if their ideal points come from the same probability distribution and are *statistically independent* of each other. Note, however, that in this case there should be little objection to redrawing constituency boundaries. Obviously, the problem of maximizing total expected utility could then be readily resolved by creating constituencies of equal population size and giving each representative the same weight if the number of constituencies is fixed—or by creating an all-encompassing, single constituency if not. This observation motivates the consideration of citizens that are a priori identical in a weaker sense: their ideal points come from the same probability distribution but are *positively correlated* within the constituencies. For this scenario, a *degressively proportional rule* remains optimal for the quadratic utility specification, but the right degree of degressivity depends on the given vector of population sizes. And, importantly, total expected utility is maximized by a *linear rule* if voters are risk-neutral, i.e., if utility falls linearly in distance, and the degree of within-constituency similarity (or dissimilarity between constituencies) is sufficiently high.

The design of welfare-maximizing voting rules for two-tier systems of representative democracy has received formal mathematical consideration only quite recently.[2] Barberà and Jackson (2006) study the design of efficient voting rules in a fairly general setup for binary decisions. They derive a square root allocation rule for the so-called "fixed-size-of-blocks model", which assumes a great degree of independence between the preferences of members of the same constituency. By contrast, they show a directly proportional allocation of weights to be optimal in their "fixed-number-of-blocks model", which reflects strong preference alignments between individuals within the same constituency (and independence across constituencies). These results are corroborated by Beisbart and Bovens (2007). Closely related, Beisbart et al. (2005) evaluate total expected utility under different decision

[2]Historically, most attention has been devoted to giving each citizen an equally effective voice in elections (cf. *Reynolds v. Sims*, 377 U.S. 533, 1964). In two-tier voting systems, this calls for an a priori equal chance of each voter to indirectly determine the policy outcome. For binary policy spaces, Penrose (1946) has shown that individual powers are approximately equalized if voting weights of the representatives are chosen such that their Penrose–Banzhaf voting powers (Penrose 1946; Banzhaf 1965) are proportional to the square root of the corresponding population sizes. An extension to convex policy spaces is provided by Maaser and Napel (2007) and Kurz et al. (2014).

rules for the Council of Ministers of the European Union and the premise that proposals always affect all individuals from a given country identically. Koriyama et al. (2013) argue in great generality that a utilitarian ideal requires vote allocation rules to be degressively proportional.

The considered objective of maximum total utility is intimately linked with achieving congruence between individual preferences and the collective policy. For binary decisions that are taken by the citizens directly (corresponding to the degenerate case of singleton constituencies and uniform weights), Rae (1969) has shown that the probability that the average citizen "has his way" (i.e., is in agreement with the voting outcome) is maximized by 50 % majority rule.[3] But the outcome of indirect, two-tier decision processes can easily deviate from that of direct democracy: even under simple majority rule it is possible that the alternative adopted by the body of representatives is supported only by a minority of all citizens.

The degree of majoritarianism of a two-tier system decreases in the expected difference between the size of the popular majority camp and the number of citizens in favor of the assembly's decision. Felsenthal and Machover (1998, pp. 63–78; 1999) study this so-called *mean majority deficit* in a binary voting model. They find it to be minimal under a square root allocation of voting weights.[4] As shown by Felsenthal and Machover, minimization of the mean majority deficit can also be interpreted in a somewhat utilitarian vein, namely as maximizing the sum of citizens' indirect voting power as measured by the non-normalized Penrose–Banzhaf index. Kirsch (2007) considers optimal weights for a related notion of majoritarian deficit. Similarly, Feix et al. (2008) investigate the *probability* of situations where the decision taken by the representatives and a hypothetical referendum decision diverge.[5] All these investigations consider the case of binary alternatives. Moving to richer policy spaces, Maaser and Napel (2012) analyze the expected discrepancy between a two-tier and a direct-democratic single-tier system in a one-dimensional spatial voting model.

A dichotomous pattern has emerged from this literature: rules that relate voting weights to the square root of population sizes have been found to be optimal under various objective functions if citizens are assumed to be homogeneous in the sense of having *independent and identically distributed* (i.i.d.) preferences. But square root rules cease to be optimal, and often a linear rule replaces them, if dependence of some sort or another is introduced. Investigations that highlight the critical role played by the degree of similarity within constituencies as opposed to that between constituencies include Gelman et al. (2002), Barberà and Jackson (2006), Kirsch (2007), Beisbart and Bovens (2007), Feix et al. (2008), Kaniovski

[3]Dubey and Shapley (1979) provide a generalization of this result to the domain of all simple games.

[4]Felsenthal and Machover refer to this allocation rule as the *second square root rule* in order to distinguish it from Penrose's (1946) (first) square root rule, which requires representatives' *voting powers*—rather than their weights—to be proportional to the square roots of their constituencies' population sizes.

[5]This situation is known in the social choice literature as a *referendum paradox* (see, e.g., Nurmi 1998).

(2008), and Maaser and Napel (2012).[6] For example, extending the main result of Felsenthal and Machover (1999) from $\{0, 1\}$-choices to the convex policy space $[0, 1]$, Maaser and Napel (2012) find that the *direct democracy deficit* is minimized when voting weights are allocated to representatives in proportion to the square root of constituency population sizes if ideal points are i.i.d. However, if sufficiently strong positive correlation of preferences within each constituency is introduced, then the best weight allocation rule is linear instead.

2 Model

We will consider a different objective function here than in Maaser and Napel (2012). But the baseline model of two-tier decision making is the same in both papers. The following description and overlapping parts of the analysis will draw directly on the presentation in Maaser and Napel (2012).

Consider the partition $\mathfrak{C} = \{\mathcal{C}_1, \ldots, \mathcal{C}_r\}$ of a large voter population into r *constituencies* with $n_j = |\mathcal{C}_j| > 0$ members each. Let $n \equiv \sum_j n_j$ and all n_j be odd numbers for simplicity. The preferences of any voter $i \in \{1, \ldots, n\} = \bigcup_j \mathcal{C}_j$ are assumed to be single-peaked with *ideal point* ν^i in a convex one-dimensional *policy space* $X \subset \mathbb{R}$, i.e., a finite or infinite interval. These ideal points are conceived of as realizations of random variables with an identical continuous a priori distribution; any given profile $(\nu^1, \ldots, \nu^n)$ of ideal points is interpreted as reflecting voter preferences on a specific one-dimensional policy issue (a tax level, expenditure on a public good, extent of redistribution, boldness of pension reform, etc.).

A collective decision $x \in X$ on the issue at hand is taken by an assembly or council of representatives $\mathcal{R}$ which consists of one representative from each constituency. Without going into details, we assume that the preferences of $\mathcal{C}_j$'s representative are congruent with its median voter, i.e., representative j has ideal point

$$\lambda_j = \text{median}\{\nu^i : i \in \mathcal{C}_j\}.$$

This is clearly an idealizing abstraction because political agents can often exploit informational asymmetries in order to pursue their own rather than their principal's preferences (e.g., concerning their privileges—see Gerber and Lewis 2004 for empirical evidence on the effect of constituency heterogeneity on the alignment between representative and median voter).

In the top-tier assembly $\mathcal{R}$, each constituency $\mathcal{C}_j$ has *voting weight* $w_j \geq 0$. Any subset $S \subseteq \{1, \ldots, r\}$ of representatives which achieves a combined weight $\sum_{j \in S} w_j$ above $q \equiv 0.5 \sum_{j=1}^{r} w_j$, i.e., comprises a *simple majority* of total weight,

[6] Also see Felsenthal and Machover (1998, pp. 70ff).

can implement a policy $x \in X$. So collective decisions are taken according to the weighted voting rule $[q; w_1, \ldots, w_r]$.

Let $\lambda_{k:r}$ denote the k-th leftmost ideal point amongst the representatives (i. e., the k-th order statistic of $\lambda_1, \ldots, \lambda_r$) and consider the random variable P defined by

$$P = \min \left\{ l \in \{1, \ldots, r\} \colon \sum_{k=1}^{l} w_{k:r} > q \right\}.$$

For a generic weight vector $(w_1, \ldots, w_r)$, representative $P:r$'s ideal point, $\lambda_{P:r}$, is the unique policy that beats any alternative $x \in X$ in a pairwise vote, i. e., it constitutes the core of the voting game in $\mathcal{R}$ with weights $w_1, \ldots, w_r$ and a 50 % quota. Without any formal analysis of decision procedures that might be applied in $\mathcal{R}$ (see Banks and Duggan 2000, or Cho and Duggan 2009), we assume that the policy agreed in the council coincides with the ideal point of *pivotal representative* $P:r$. In summary, the policy outcome produced by the two-tiered voting system is

$$x_{\mathcal{R}} = \lambda_{P:r} \, .$$

For ideal point profile $(\nu^1, \ldots, \nu^n)$ the total utility that the society receives from $x_{\mathcal{R}}$ is

$$\bar{U} = \sum_{i=1}^{n} -|\nu^i - x_{\mathcal{R}}|, \quad \text{or} \tag{1a}$$

$$\hat{U} = \sum_{i=1}^{n} -(\nu^i - x_{\mathcal{R}})^2 \tag{1b}$$

if for each voter utility decreases (a) linearly or (b) quadratically in the distance between his ideal point and the outcome.

Taking partition $\mathfrak{C}$ as given we would like to answer the following question: *Which allocation of voting weights maximizes the total expected utility of the two-tier voting system?* Or, more formally, we search for weight allocation rules W which approximately solve the problems

$$\max_{w_1,\ldots,w_r} \mathbf{E}\left[\bar{U}\right], \quad \text{and} \tag{2a}$$

$$\max_{w_1,\ldots,w_r} \mathbf{E}\left[\hat{U}\right], \tag{2b}$$

respectively, where by an "allocation rule" we mean a simple mapping W which assigns weights $(w_1, \ldots, w_r) = W(\mathcal{C}_1, \ldots, \mathcal{C}_r)$ to any given partition of a large population. Our criterion for acceptably "simple" mappings $W \colon \mathfrak{C} \mapsto (w_1, \ldots, w_r)$ will be that they are *power laws*, i.e., $w_j = n_j^{\alpha}$ for some constant $\alpha \in [0, 1]$.

This class of mappings nests the square root and linear rules which have played prominent roles in the previous literature.[7]

3 Analysis

Under the model's assumptions, it can be shown that societal welfare $\bar{U}$ would be maximized if, for any realization of voter preferences, we had

$$x_{\mathcal{R}} = \text{median}\{\nu^1, \ldots, \nu^n\}.$$

That is, it is the unconstrained ideal to choose the preferred policy of the median individual in the union (see, e.g., Schwertman et al. 1990). This policy outcome would also be brought about by frictionless collective decision-making in a full assembly of all citizens under simple majority rule since it beats every alternative policy in a pairwise vote. Because of the loss of information that results from only aggregating the votes of the top-tier representatives, however, $x_{\mathcal{R}}$ does generally *not* coincide with the median ideal point in the population. Problem (2a) is thus equivalent to that of minimizing the expected value of $|x_{\mathcal{R}} - \text{median}\{\nu^1, \ldots, \nu^n\}|$, which is referred to as the *direct democracy deficit* in Maaser and Napel (2012).[8]

While the median has the property of minimizing the sum of absolute distances, the sum of squared distances is minimized by the mean (see, e.g., Cramér 1946, Sect. 15.4). Thus, the ideal non-voting solution to problem (2b) would be to always implement the policy that corresponds to the *mean* of ideal points $\{\nu^1, \ldots, \nu^n\}$. Our maximization problem can, therefore, be reframed in the case of quadratic utility functions as follows: by which simple weight allocation rule do we achieve a particularly "small" expected distance between $x_{\mathcal{R}}$ and the mean voter position? In principle, an estimate of the overall mean could be obtained by taking the n_j-weighted mean of $\lambda_1, \ldots, \lambda_r$. If, however, representatives' positions are aggregated by voting under strategic interaction rather than being averaged (e.g., by a bureaucrat) then the outcome $x_{\mathcal{R}}$ at the top-tier will match one of the representatives' positions in the considered spatial voting model, namely their n_j^α*-weighted median* in our model. This will usually differ from the n_j*-weighted mean*. Optimal statistical aggregation *by averaging* does not really help in solving the problem of optimal aggregation *by voting*.

---

[7]To be precise, Penrose's square root rule is nested only asymptotically, namely when $\mathfrak{C}$ involves a great number r of constituencies with a regular size distribution. See Lindner and Machover (2004) and Chang et al. (2006) on the vanishing difference between voting weights and voting powers as $r \to \infty$.

[8]Note that even though total utility from the decisions which result from the considered two-tier process typically falls short of the global maximum achieved under a direct democracy, representative democracy has a number of advantages. These presumably also generate utility for citizens which is not considered in our model.

If the ideal points of voters $i \in \mathcal{C}_j$ are pairwise independent and come from an arbitrary identical distribution F with positive density f on X, then its median position λ_j asymptotically has a normal distribution with mean $\mu = F^{-1}(0.5)$ and standard deviation

$$\sigma_j = \frac{1}{2\, f(\mu)\sqrt{n_j}} \tag{3}$$

(see, e.g., Arnold et al. 1992, p. 223). The variance of the position of $\mathcal{C}_j$'s representative is the smaller, the greater the population size n_j.

This implies that even in the seemingly trivial case of uniform weights $w_1 = \ldots = w_r$, the top-tier decision $x_{\mathcal{R}} \in X$ has a rather non-trivial distribution when constituency sizes differ. Namely, $x_{\mathcal{R}}$ is then an order statistic of *differently* distributed random variables, for which relatively few limit results are known. For non-identical weights $w_1, \ldots, w_r$, $x_{\mathcal{R}}$ is a combinatorial function of such order statistics. Therefore, it seems extremely hard—at least to us—to obtain or approximate solutions to (2a) and (2b) analytically. We will now briefly look at two special cases in order to develop some intuition, and then turn to computer simulations in Sect. 4.

First, consider the trivial case of equipopulous constituencies. Any uniform weight allocation $w_1 = \ldots = w_r > 0$ then maximizes total expected utility and the optimal value of α remains undetermined. Under identical weights, the pivotal representative's ideal point is the (unweighted) median of $\lambda_1, \ldots, \lambda_r$. How close this comes to the population's sample median and mean, respectively, will depend on the number of symmetric constituencies in the partition.[9]

So far no assumptions have been made regarding how voter ideal points are jointly distributed. It can be convincingly argued that—for the kind of constitutional design problem that we are dealing with—specific knowledge about individual preferences should be ignored. From behind the constitutional "veil of ignorance" all citizens should be considered *identical* a priori. This corresponds to drawing every ideal point ν^i from the *same* marginal probability distribution F. However, such a constitutional a priori perspective does not necessarily entail that preferences of citizens must also be conceived of as *independent* of each other. It is true that the i.i.d. assumption for all ideal points ν^i with $i \in \bigcup_j \mathcal{C}_j$, i.e., consideration of the product distribution F^n, is a particularly compelling benchmark. Still, the partition $\mathfrak{C}$ may have reasons that need to be acknowledged behind the "veil of ignorance" (e.g., geographic barriers, ethnics, language, or religion). For these reasons voter preferences are likely to be more closely connected within constituencies than across them.

As a second case of interest, suppose that $\nu^i = \nu^h$ whenever $i, h \in \mathcal{C}_j$. This carries the notion that citizens have on average closer links with each other within

[9] See Beisbart and Bovens (2013) for a related investigation in a binary voting model. They ask the worst-case question: which *number* of equipopulous districts *maximizes* the mean majority deficit?

constituencies than across constituencies to its extreme. Problem (2a) has a clear-cut solution in this situation: $\mathbf{E}[\bar{U}]$ is maximal if the linear weight allocation rule $w_j = n_j$ for $j = 1, \ldots, r$, i.e., $\alpha^* = 1$, is used. Perfect correlation within constituencies implies that the ordered ideal points of all citizens $i = 1, \ldots, n$,

$$\nu^{1:n} \leq \nu^{2:n} \leq \nu^{3:n} \leq \ldots \leq \nu^{n-1:n} \leq \nu^{n:n},$$

can be written as

$$\underbrace{\lambda_{1:r} = \ldots = \lambda_{1:r}}_{n_{1:r}\text{times}} \leq \underbrace{\lambda_{2:r} = \ldots = \lambda_{2:r}}_{n_{2:r}\text{times}} \leq \ldots \leq \underbrace{\lambda_{r:r} = \ldots = \lambda_{r:r}}_{n_{r:r}\text{times}}.$$

Thus, weights proportional to population sizes make representative j pivotal in $\mathcal{R}$ if and only if his policy position (and thus that of all $\mathcal{C}_j$-citizens) is also the population median. In the non-degenerate case of high but not perfect correlation within constituencies this optimality of proportional weights can be expected to apply *approximately*. The simulations reported in Sect. 4 indeed confirm this intuition: with linear individual utility functions total expected utility is maximized by an essentially linear rule provided that the ideal points of the citizens vary noticeably more across than within constituencies.

The above extreme case is also instructive to appreciate that a linear rule cannot be optimal in general when individual utility decreases *quadratically* in the distance between ideal point and outcome. When constituencies differ in population size, the overall frequency distribution of policy positions will typically not be symmetric. It will be skewed to the right if a majority of the large constituencies prefers a policy to the left of the center (see Fig. 1), and it is skewed to the left if the large constituencies have ideal points on the right. But then the population median (which would result from $\alpha = 1$) does not provide a good estimate of the population mean: the sample median is necessarily located to the left of the sample mean if the distribution is skewed to the right (just like average income is larger than median income if there are many small incomes and a few very large ones). A value of $\alpha < 1$ then produces a smaller deviation between the pivotal representative's ideal point and the mean ideal point. Figure 1 illustrates this by taking EU27 members as an example. In the figure all citizens within each constituency have identical policy preferences drawn from a uniform distribution on $[-1, 1]$. For the depicted right-skewed realization $(\nu_1, \ldots, \nu_n)$ of ideal points, $\alpha = 0.71$ is best among the considered parameters $\alpha \in \{0, 0.01, \ldots, 1\}$: the associated outcome $x_{\mathcal{R}}$ is as close as possible to the mean of all ideal points. The same degressivity parameter $\alpha = 0.71$ would be optimal for the ideal point realization $(\nu'_1, \ldots, \nu'_n)$ with $\nu'_i = -\nu_i$ for all $i = 1, \ldots, n$, which is skewed to the left. For realizations that give rise to an essentially symmetric frequency distribution, $\alpha = 0.71$ performs as well as any alternative value (such as $\alpha = 1$). We can, therefore, conclude that $\alpha = 1$ must be suboptimal if one averages over all possible frequency distributions, i.e., considers the expected value $\mathbf{E}[\hat{U}]$. The *optimal* value of α depends on the constituency configuration at hand as well as

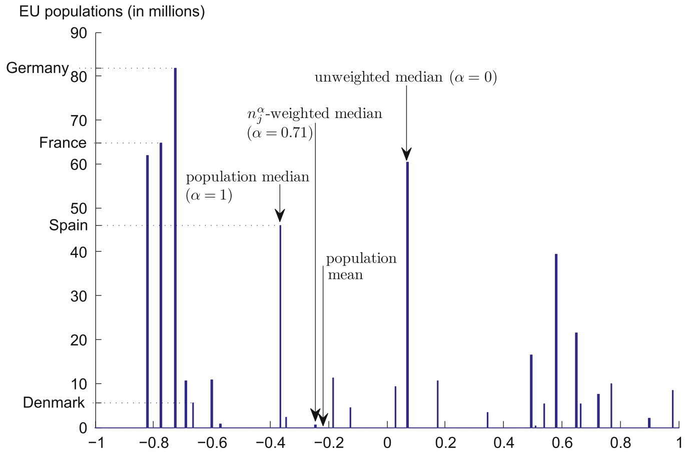


Fig. 1 Sample frequency distribution of policy positions for EU27 member countries (2010 Eurostat population data)

on the theoretical distribution of individual ideal points.[10] It might but need not be close to 0.5.

When we consider non-degenerate degrees of correlation between the ideal points within a given constituency, it is even more difficult to come up with a clear intuition for what the best degree of degressivity should be. In the benchmark case of ideal points that are all pairwise independent and drawn from the same symmetric distribution, computation of the n_j-weighted mean of $\lambda_1, \ldots, \lambda_r$ would be the theoretically best way to estimate both the location of the sample median and the sample mean. The n_j-weighted mean is sensitive to outliers amongst the representatives' ideal points. This rules out optimality of $\alpha = 0$ because uniform weights select the median representative's ideal point and hence disregard any information about outliers. But a too great value of α would enable representatives from large constituencies to implement their preferred policy even if they happen to be outliers. It is not obvious at the outset what "too great" means and which α strikes the right balance.

An admittedly crude intuitive argument in favor of $\alpha = 0.5$ runs as follows. First consider the linear utility specification, so that the theoretical ideal is to approximate the population's *median voter* as well as possible. If all voter ideal points ν^i are i.i.d. then each individual $i = 1, \ldots, n$ a priori has probability $1/n$ to be the population median. The latter is hence located in constituency $\mathcal{C}_j$ with probability n_j/n. This makes weights which induce top-tier pivot probabilities proportional to the respective population sizes a particularly reasonable starting point. As Kurz et al.

[10] The problem of finding the optimal value of α bears some resemblance to choosing an appropriate power-law transformation in order to improve the symmetry of a skewed empirical distribution (see, e.g., Yeo and Johnson 2000).

(2014) have shown, proportionality between the probability of the event $\{j = P : r\}$ and n_j can be achieved approximately by selecting weights w_j that are proportional to the *square root* of n_j in the i.i.d. case.

Second, for the quadratic utility specification, the goal is to approximate the population's *mean voter* by selecting a particular weighted median of the representatives. The mean voter is a virtual one who does not belong to any particular constituency. Notably, the mean ideal point will almost surely differ from those of *all* voters $i \in \{1, \dots, n\}$ when the ideal point distribution has a density. Thus, the intuition provided for a square root rule in the case of linear utility does not apply directly to the case of quadratic utility functions. However, for the symmetric ideal point distributions which we focus on in this paper, the population mean and median will be very close to each other if all voters are pairwise independent. One may conjecture, therefore, that $\alpha = 0.5$ will work well under an i.i.d. assumption irrespective of the utility specification.

4 Simulations

Since we are unable to obtain more precise analytical insights—let alone any useful approximation of $\mathbf{E}[\bar{U}]$ or $\mathbf{E}[\hat{U}]$ as a function of α—we apply the *Monte-Carlo method*. It exploits that the empirical average of s independent realizations of $\bar{U} = \sum_{i=1}^{n} -|\nu^i - x_{\mathcal{R}}|$ and $\hat{U} = \sum_{i=1}^{n} -(\nu^i - x_{\mathcal{R}})^2$ converges to $\mathbf{E}[\bar{U}]$ and $\mathbf{E}[\hat{U}]$, respectively, as $s \to \infty$ by the law of large numbers.

In order to obtain realizations of $\bar{U}$ and $\hat{U}$ for the case of i.i.d. voter ideal points, we first draw n (pseudo-)random numbers from a given distribution F, giving rise to a list $\mathbf{v} = (\nu^1, \dots, \nu^n)$.[11] Second, $\mathbf{v}$ is sorted within consecutive blocks of size $n_1, n_2, \dots, n_r$ in order to obtain the corresponding realizations of the constituency medians $\lambda_1, \lambda_2, \dots, \lambda_r$. We then infer the weighted median of these, using weights $w_j = n_j^{\alpha}$ for values of α which range from 0 to 1 in steps of 0.01, and thus obtain $x_{\mathcal{R}}$ for each value of α. The resulting values of $\bar{U}$ and $\hat{U}$ are recorded, and the procedure is repeated for one million iterations. Finally, we determine the values of α, denoted by $\bar{\alpha}^*$ and $\hat{\alpha}^*$ which produced the largest average total utility $\bar{U}$ and $\hat{U}$, respectively.

In our simulations we typically consider sets of $r = 25$ constituencies. Experience suggests that simulation results then do no longer exhibit strong dependence on the combinatorial peculiarities of the configuration at hand (this would be the case for significantly smaller numbers of constituencies). Most of the

---

[11] Since the considered number of voters in each constituency $\mathcal{C}_j$ is large ($n_j \gg 50$), the respective population and constituency medians will approximately have normal distributions irrespective of the specific F which one considers. For the sake of completeness, let it still be mentioned that individual ideal points were drawn from a standard uniform distribution $\mathbf{U}(0, 1)$ in our simulations. The MATLAB source code is available upon e-mail request.

Table 1 Welfare-maximal α for i.i.d. voters

(a) *Linear utility*

	$\bar{\alpha}^*$				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
$\mathbf{U}(1000, 3000)$	**0.52**	**0.51**	**0.51**	**0.51**	**0.53**
$\mathbf{N}(2000, 200)$	**0.43**	**0.62**	**0.65**	**0.57**	**0.44**
$\mathbf{N}(2000, 400)$	**0.48**	**0.52**	**0.50**	**0.47**	**0.51**
$\mathbf{P}(1.0, 200)$	**0.52**	**0.52**	**0.52**	**0.53**	**0.52**

(b) *Quadratic utility*

	$\hat{\alpha}^*$				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
$\mathbf{U}(1000, 3000)$	**0.51**	**0.51**	**0.48**	**0.51**	**0.53**
$\mathbf{N}(2000, 200)$	**0.49**	**0.62**	**0.65**	**0.57**	**0.44**
$\mathbf{N}(2000, 400)$	**0.52**	**0.52**	**0.50**	**0.46**	**0.53**
$\mathbf{P}(1.0, 200)$	**0.52**	**0.52**	**0.51**	**0.53**	**0.52**

considered population configurations are artificial: sizes $n_1, \ldots, n_r$ are obtained by drawing random numbers from a specified distribution. The entry $\mathbf{U}(10^3, 3 \cdot 10^3)$ in Table 1, for instance, indicates that realizations of $\mathfrak{C}$ are considered for which each constituency size between 1,000 and 3,000 voters had *uniform* probability. Besides the uniform distribution, also truncated *normal* distributions $\mathbf{N}(\mu, \sigma)$ and *Pareto* distributions $\mathbf{P}(\kappa, \theta)$ with skewness parameter κ and threshold parameter θ have been employed in order to generate population configurations. For each "distribution type" of the population configuration, five independent realizations of $n_1, \ldots, n_r$ have been investigated. So Table 1 reports the respective optimal values $\bar{\alpha}^*$ (linear utility) and $\hat{\alpha}^*$ (quadratic utility) for altogether 20 different configurations.

The 95 %-confidence intervals around the empirical mean of $\bar{U}$ and $\hat{U}$ are typically too wide to rule out that a neighbor of the reported best value of α produces a higher level of welfare. However, differences are significant when sufficiently distinct values like $\alpha = 0.5$ and $\alpha = 1$ are compared.[12] The obtained estimates of $\mathbf{E}[\Delta]$ are in most cases unimodal functions of α, i.e., increasing on $[0, \alpha^*)$ and decreasing on $(\alpha^*, 1]$. Overall, results in Table 1 are suggesting strongly that a square root allocation rule is close to being optimal (within the class of elementary power laws) if the ideal points of all voters are independent and identically distributed.

Concerning cases in which the ideal points of citizens are *not* independent and identically distributed, we focus on a special type of positive correlation within constituencies. In particular, we determine individual ideal points ν^i by a two-step random experiment: first, we draw a constituency-specific parameter μ_j independently for each $j = 1, \ldots, r$ from an identical distribution G with standard

[12]In particular, variation in population sizes $n_j \sim \mathbf{N}(2000, 200)$ is rather small. This results in an objective function that is essentially flat for a large range of values of α.

Table 2 Welfare-maximal α for two different preference dissimilarity ratios d

(a) *Linear utility*

		$\bar{\alpha}^*$				
		(1)	**(2)**	**(3)**	**(4)**	**(5)**
$\mathbf{U}(1000, 3000)$	$d = 8$	**0.96**	**0.96**	**0.96**	**0.97**	**0.96**
	$d = \infty$	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
$\mathbf{N}(2000, 200)$	$d = 8$	**0.96**	**0.95**	**0.97**	**0.96**	**0.97**
	$d = \infty$	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
$\mathbf{N}(2000, 400)$	$d = 8$	**0.96**	**0.96**	**0.95**	**0.97**	**0.95**
	$d = \infty$	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
$\mathbf{P}(1.0, 200)$	$d = 8$	**0.97**	**0.96**	**0.96**	**0.97**	**0.97**
	$d = \infty$	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**

(b) *Quadratic utility*

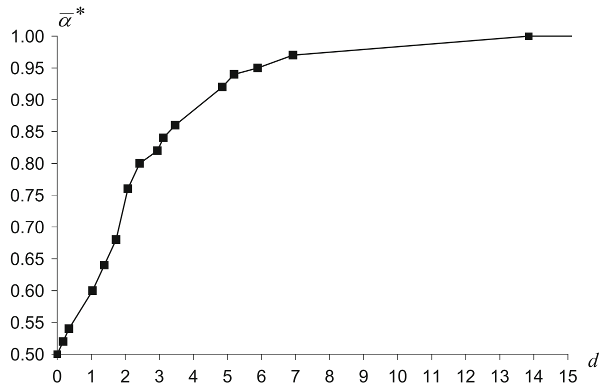
		$\hat{\alpha}^*$				
		(1)	**(2)**	**(3)**	**(4)**	**(5)**
$\mathbf{U}(1000, 3000)$	$d = 8$	**0.48**	**0.49**	**0.52**	**0.53**	**0.48**
	$d = \infty$	**0.51**	**0.49**	**0.54**	**0.55**	**0.50**
$\mathbf{N}(2000, 200)$	$d = 8$	**0.47**	**0.51**	**0.65**	**0.53**	**0.62**
	$d = \infty$	**0.53**	**0.54**	**0.56**	**0.61**	**0.51**
$\mathbf{N}(2000, 400)$	$d = 8$	**0.50**	**0.54**	**0.47**	**0.54**	**0.52**
	$d = \infty$	**0.49**	**0.51**	**0.51**	**0.51**	**0.47**
$\mathbf{P}(1.0, 200)$	$d = 8$	**0.64**	**0.56**	**0.55**	**0.68**	**0.66**
	$d = \infty$	**0.64**	**0.56**	**0.55**	**0.68**	**0.66**

deviation σ_{ext}. Parameter σ_{ext} captures the degree of *external heterogeneity* between $\mathcal{C}_1, \ldots, \mathcal{C}_r$ for the policy issue at hand. The realization of parameter μ_j is taken to reflect the expected ideal point of citizens from $\mathcal{C}_j$ on a given policy issue. Each citizen $i \in \mathcal{C}_j$ is then assigned an individual ideal point ν^i from a distribution F_{μ_j} which has mean μ_j and is otherwise just a shifted version of some distribution $F \equiv F_0$ for each constituency $j = 1, \ldots, r$.[13] F's standard deviation σ_{int} is a measure of the *internal heterogeneity* in any constituency. It reflects opinion differences within any given $\mathcal{C}_j$. In summary, our second set of simulations has taken the ideal points of all citizens to be *identically distributed* with convolved a priori distribution $G * F$, but to involve *dependencies*: citizens in constituency $\mathcal{C}_j$ all experience the same shift μ_j, which is independent of μ_k for any $k \neq j$.

The ratio $\sigma_{ext}/\sigma_{int} =: d$ between external and internal heterogeneity provides a measure of the degree to which citizens are more similar within than between constituencies or, loosely speaking, the *preference dissimilarity* of the constituencies. In the i.i.d. case *no* dissimilarity exists between different constituencies, i.e., results in Table 1 are based on $d = 0$. Table 2 reports optimal values $\bar{\alpha}^*$ and $\hat{\alpha}^*$ for the same configurations as in Table 1 and two positive dissimilarity levels, namely $d = 8$

[13]Specifically, we draw μ_j from a uniform distribution $\mathbf{U}(-a, a)$ with variance σ^2_{ext}, and then obtain $\nu^i = \mu_j + \varepsilon$ with $\varepsilon \sim \mathbf{U}(0, 1)$.

Fig. 2 Welfare-maximal α for EU27 and linear utility as dissimilarity ratio d is varied



and the degenerate case of infinite dissimilarity ($\sigma_{int} = 0$). While results for i.i.d. ideal points did not significantly differ between the linear specification of individual utility functions and the quadratic one in Table 1, this is no longer the case when significant preference correlations exist.

The optimality of $\bar{\alpha}^* = 1$ as $d \to \infty$ for the linear specification has already been explained in our theoretical discussion in Sect. 3 (considering fixed $\sigma_{ext} > 0$ and $\sigma_{int} \to 0$). The findings reported in Table 2a indicate that this result extends in close approximation to more moderate levels of dissimilarity such as $d \geq 8$. Figure 2 demonstrates that a situation in which nearly linear voting weight allocations maximize $\mathbf{E}[\bar{U}]$ arises quickly as the preference dissimilarity which underlies the policy ideals of representatives in $\mathcal{R}$ increases. The figure considers $r = 27$ and a population configuration based on recent Eurostat data for members of the European Union.[14] The EU Council of Ministers is the predominant example of a two-tier voting system because its members officially represent national governments and, eventually, the citizenries of the member states. Note, however, that the current weighted voting rules for the Council, and also its future ones as codified in the Treaty of Lisbon, involve supermajority requirements in multiple dimensions, while Fig. 2 is based on the assumption of a 50 % decision quota. We leave an investigation of the effect of supermajority rules on the maximizer (and maximum) of utilitarian welfare in our spatial voting framework to future research.

The optimal levels of $\hat{\alpha}^*$ for a quadratic utility specification, displayed in Table 2b, fail to show convergence to any specific rule as $d \to \infty$. In particular, it does not seem to make a significant difference whether dissimilarity is moderate or extreme. Moreover, the reported values of $\hat{\alpha}^*$ do not differ noticeably from their i.i.d. counterparts in Table 1 except for Pareto-distributed population configurations (where constituency sizes have a skewed distribution).

[14]We have used 2010 population data measured in 1,000 individuals for computational reasons. This corresponds with the "block model" in Barberà and Jackson (2006), which supposes that a constituency can be subdivided into equally sized "blocks" whose members have perfectly correlated preferences within blocks, but are independent across blocks.

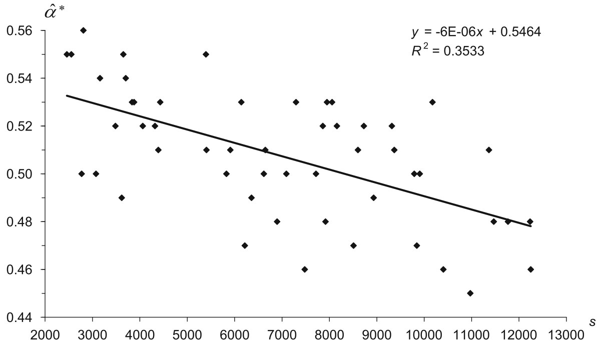


Fig. 3 Optimal $\hat{\alpha}$ for 52 population configurations with sample standard deviations s

We argued in our discussion of Fig. 1 that $\alpha = 1$ should not be expected to be optimal when individual utility functions are quadratic and $d \to \infty$, and that it is not clear which particular $\hat{\alpha}$ should be optimal. Table 2 suggests vaguely that a square root allocation might actually do best when constituency sizes are drawn from a symmetric distribution (uniform or normal). But certainly more weight needs to be given to large constituencies than under a square root law when the population distribution is skewed (Pareto).

Note that even if the distribution of *population sizes* $n_1, \ldots, n_r$ is symmetric, the realized frequency distributions of *ideal points* will be skewed more often than not. For instance, a frequency distribution like the one displayed in Fig. 1 will still be common even if we have constituency sizes that range equidistantly from some smallest value $\underline{n}$ to a largest value $\overline{n}$ (mimicking a uniform distribution on $[\underline{n}, \overline{n}]$). So some degressively proportional weighting scheme raises total expected utility relative to a linear rule. We conjecture that, for symmetric distributions of constituency sizes, the average distance between the sample median and the sample mean is larger, the larger the variance of $n_1, \ldots, n_r$. Therefore, the greater the variance of $n_1, \ldots, n_r$, the smaller the optimal value $\hat{\alpha}^*$. This hypothesis is supported by our simulation data. In particular, Fig. 3 displays the welfare-maximizing level $\hat{\alpha}^*$ in case of the quadratic utility specification and degenerate preference dissimilarity ($d = \infty$) for altogether 52 distinct population configurations that were drawn either from uniform and (truncated) normal distributions with $r = 25$ or $r = 35$. A higher standard deviation s of the population sizes $n_1, \ldots, n_r$ visibly translates into a smaller optimal value $\hat{\alpha}^*$. The slope of the corresponding regression line is not very steep, but it is significantly different from zero. Still, a value of $\alpha = 0.5$ is never very far off. This is in line with the findings in Table 2 for the symmetric distributions of constituency sizes.

5 Concluding Remarks

The findings of our investigation of utilitarian welfare or total expected utility of the citizens in a spatial voting model might be summarized—cum grano salis—as supporting the conclusions of the related literature on binary voting models (see Sect. 1). In particular, *if* the preferences of the voters are characterized by independent and identically distributed (i.i.d.) ideal points over a one-dimensional policy space, then using a *square root rule* for allocating voting weights performs best. This is irrespective of whether voters' utility decreases linearly or quadratically in distance from their policy ideal (corresponding to risk neutrality or a particular extent of risk aversion when facing uncertain collective decisions). Unfortunately, we could provide but a vague intuition for *why* a square root law obtains.

Our findings are also consistent with the binary voting literature in that optimality of a square root rule—be it elementary like $w_j = n_j^{0.5}$ or sophisticated like the seminal suggestion by Penrose (1946)—tends *not* to extend to situations in which the i.i.d. assumption is violated. It has increasingly come to be understood that when voters have a priori identical random preferences in the binary case or on some richer space, like the one considered here, and these preferences exhibit *positive correlation* within constituencies, then there is a potentially very rapid phase transition from $\alpha = 0.5$ to $\alpha = 1$ performing best.

However, the results shown in Table 2b and Fig. 3 cast some doubt on this dichotomy between using a square root rule for similar constituencies and a linear rule for sufficiently dissimilar constituencies. Even though $\alpha = 0.5$ ceases to be welfare-maximizing in the considered class of allocation rules, especially when the distribution of population sizes is skewed, it tends to perform better than a linear rule when individuals have a quadratic utility function. This is surprising given that our assumptions such as single-peakedness of preferences in a one-dimensional policy space and the prominent role of the (weighted) median are rather straightforward generalizations from the realm of binary voting.

Note additionally that our utilitarian welfare investigation builds on the restrictive postulate that different individuals derive the same satisfaction or dissatisfaction when a policy at a certain distance from their ideal point is implemented. In other words, we conduct interpersonal comparisons of utility. These cannot be avoided by any utilitarian welfare analysis in economics or political science. And, here, they can be defended by the a prioristic nature of the investigation: they express the value judgment that all individuals should be treated as anonymous equals in constitutional analysis. Still, the fact that our findings differ for different specifications of voter utility—and rather distinct conclusions might be derived concerning the most desirable allocation of voting weights in, e.g., the EU's Council of Ministers—might be seen as weakening the appeal of total expected utility as a guide to the "best" weight allocation rule.

References

Arnold, B. C., Balakrishnan, N., & Nagaraja, H. N. (1992). *A first course in order statistics*. New York: Wiley.

Banks, J. S., & Duggan, J. (2000). A bargaining model of collective choice. *American Political Science Review*, *94*(1), 73–88.

Banzhaf, J. F. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review*, *19*(2), 317–343.

Barberà, S., & Jackson, M. O. (2006). On the weights of nations: Assigning voting weights in a heterogeneous union. *Journal of Political Economy*, *114*(2), 317–339.

Beisbart, C., & Bovens, L. (2007). Welfarist evaluations of decision rules for boards of representatives. *Social Choice and Welfare*, *29*(4), 581–608.

Beisbart, C., & Bovens, L. (2013). Minimizing the threat of a positive majority deficit in two-tier voting systems with equipopulous units. *Public Choice*, *145*(1–2), 75–94. doi:10.1007/s11127-011-9810-2.

Beisbart, C., Bovens, L., & Hartmann, S. (2005). A utilitarian assessment of alternative decision rules in the Council of Ministers. *European Union Politics*, *6*(4), 395–419.

Chang, P.-L., Chua, V. C., & Machover, M. (2006). L S Penrose's limit theorem: Tests by simulation. *Mathematical Social Sciences*, *51*(1), 90–106.

Cho, S., & Duggan, J. (2009). Bargaining foundations of the median voter theorem. *Journal of Economic Theory*, *144*(2), 851–868.

Cramér, H. (1946). *Mathematical methods of statistics*. Princeton landmarks in mathematics. Princeton: Princeton University Press [Reprinted 1999].

Dubey, P., & Shapley, L. (1979). Mathematical properties of the Banzhaf power index. *Mathematics of Operations Research*, *4*(2), 99–131.

Feix, M. R., Lepelley, D., Merlin, V., Rouet, J.-L., & Vidu, L. (2008). *Majority efficient representation of the citizens in a federal union*. Mimeo: Université de la Réunion, Université de Caen, and Université d'Orléans.

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

Felsenthal, D., & Machover, M. (1999). Minimizing the mean majority deficit: The second square-root rule. *Mathematical Social Sciences*, *37*(1), 25–37.

Gelman, A., Katz, J. N., & Tuerlinckx, F. (2002). The mathematics and statistics of voting power. *Statistical Science*, *17*(4), 420–435.

Gerber, E. R., & Lewis, J. B. (2004). Beyond the median: Voter preferences, district heterogeneity, and political representation. *Journal of Political Economy*, *112*(6), 1364–1383.

Harsanyi, J. C. (1955). Cardinal welfare, individual ethics, and interpersonal comparison of utility. *Journal of Political Economy*, *63*(4), 309–321.

Kaniovski, S. (2008). The exact bias of the Banzhaf measure of power when votes are neither equiprobable nor independent. *Social Choice and Welfare*, *31*(2), 281–300.

Kirsch, W. (2007). On Penrose's square-root law and beyond. *Homo Oeconomicus*, *24*(3/4), 357–380.

Koriyama Y., Laslier, J.-F., Macé, A. & Treibich R. (2013). *Optimal Apportionment. Journal of Political Economy*, *121*(3), 584–608.

Kurz, S., Maaser, N., & Napel, S. (2014). *On the democratic weights of nations*. Mimeo: University of Bayreuth.

Leech, D., & Leech, R. (2009). Reforming IMF and World Bank governance: In search of simplicity, transparency and democratic legitimacy in the voting rules. Warwick Economic Research Papers 914.

Lindner, I., & Machover, M. (2004). L. S. Penrose's limit theorem: Proof of some special cases. *Mathematical Social Sciences*, *47*(1), 37–49.

Maaser, N., & Napel, S. (2007). Equal representation in two-tier voting systems. *Social Choice and Welfare*, *28*(3), 401–420.

Maaser, N., & Napel, S. (2012). A note on the direct democracy deficit in two-tier voting. *Mathematical Social Sciences*, *62*(2), 174–180.

Nurmi, H. (1998). *Rational behaviour and design of institutions*. Cheltenham: Edward Elgar.

Penrose, L. S. (1946). The elementary statistics of majority voting. *Journal of the Royal Statistical Society*, *109*(1), 53–57.

Rae, D. W. (1969). Decision rules and individual values in constitutional choice. *American Political Science Review*, *63*(1), 40–56.

Schwertman, N. C., Gilks, A. J., & Cameron, J. (1990). A simple noncalculus proof that the median minimizes the sum of the absolute deviations. *American Statistician*, *44*(1), 38–39.

Yeo, I.-K., & Johnson, R. A. (2000). A new family of power transformations to improve normality or symmetry. *Biometrika*, *87*(4), 954–959.

A Priori Voting Power When One Vote Counts in Two Ways, with Application to Two Variants of the U.S. Electoral College

Nicholas R. Miller

The President of the United States is elected, not by a direct national popular vote, but by a two-tier Electoral College system in which (in almost universal practice since the 1830s) separate state popular votes are aggregated by adding up state electoral votes awarded, on a winner-take-all basis, to the plurality winner in each state. Each state has electoral votes equal in number to its total representation in Congress and since 1964 the District of Columbia has three electoral votes. At the present time, there are 435 members of the House of Representatives and 100 Senators, so the total number of electoral votes is 538, with 270 required for election (with a 269–269 tie possible). The U.S. Electoral College is therefore a *two-tier* electoral system: individual voters cast votes in the lower-tier to choose between rival slates of 'Presidential electors' pledged to one or other Presidential candidate, and the winning elector slates then cast blocs of electoral votes for the candidate to whom they are pledged in the upper tier. The Electoral College therefore generates the kind of weighted voting system that invites analysis using one of the several measures of *a priori* voting power. With such a measure, we can determine whether and how much the power of voters may vary from state to state and how individual voting power may change under different variants of the Electoral College system.

1 Individual Voting Power Under the Electoral College

Several years ago, I had a commission to write an encyclopedia entry on "Voting Power in the U.S. Electoral College" (Miller 2011), and I decided to include a chart (resembling Fig. 1) displaying individual voting power by state under the

N.R. Miller (✉)
Department of Political Science, University of Maryland Baltimore County (UMBC), 1000 Hilltop Circle, Baltimore, MD 21250, USA
e-mail: nmiller@umbc.edu

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare, DOI 10.1007/978-3-319-05158-1_11,

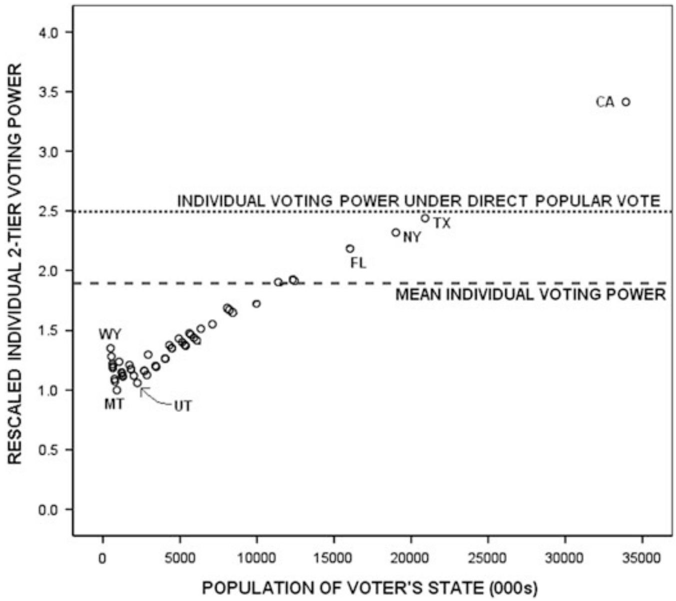


Fig. 1 Individual voting power by state population under the existing apportionment of electoral votes

apportionment of electoral votes based on the 2000 census. Having been introduced some years earlier to Dan Felsenthal and Moshe Machover's magnificent treatise on *The Measurement of Voting Power* (1998), I was confident that I had a reasonably precise understanding of the properties and proper interpretations of the various voting power measures (with which I had been broadly familiar since graduate school). I also believed that I could make the necessary calculations using the immensely useful website on *Computer Algorithms for Voting Power Analysis* created and maintained by Dennis Leech and Robert Leech.[1]

I was persuaded by Felsenthal and Machover's emphatic advice that the absolute *Banzhaf measure* is the proper measure of *a priori* voting power in the context of ordinary two-candidate or two-party elections. Given n voters, there are 2^{n-1} *bipartitions* (i.e., complementary pairs of subsets) of voters (including the pair consisting of the set of all voters and the empty set). A voter (e.g., a state) is *critical* in a bipartition if the set to which the voter belongs is *winning* (e.g., a set of states with at least 270 electoral voters) but would not be winning if the voter belonged to the complementary set. A voter's *Banzhaf score* is the total number of bipartitions in which the voter is critical. A voter's *absolute Banzhaf voting power* is the voter's Banzhaf score divided by the number of bipartitions.

[1]The website may be found at http://www.warwick.ac.uk/~ecaae/.

Felsenthal and Machover show that the absolute Banzhaf measure (unlike the 'relative' Banzhaf index or the Shapley–Shubik index) has the following directly meaningful and analytically useful probabilistic interpretation. Suppose we know nothing about individual voters except their positions with respect to the formal properties of a voting system (in this case, what state they live in) but nothing about their political inclinations, voting habits, etc., and that, from behind this 'veil of ignorance,' we wish to assess their voting power. For this purpose (though certainly not for many others), our *a priori* expectation must be that individuals vote randomly, i.e., as if they are independently flipping fair coins in what may be called a *random* (or *Bernoulli*) *election*. On this assumption, Felsenthal and Machover show that a voter's absolute Banzhaf voting power is the probability that he or she casts a *decisive* vote that determines the outcome of such a random election (e.g., that, given all other votes, breaks what would otherwise be a tie).

Now suppose likewise that we know nothing about U.S. Presidential elections other than the formal rules of the Electoral College—specifically, we know the population of each state, the total number of electoral votes, the formula for apportioning these electoral votes among the states on the basis of population, and the fact that each state's electoral votes are cast as bloc for the candidate who wins the most popular votes in the state. Absent any further information, we must assume that the total number of votes cast in a state is equal to some fixed percent of the state's apportionment population. In a two-tier voting system such as the Electoral College, voter i's *a priori* voting power is the probability that i casts a *doubly decisive* vote, i.e., one that creates or breaks what would otherwise be a tie in the popular vote in the voter's state, which in turn breaks what would otherwise be a deadlock in the Electoral College. Put otherwise, the *a priori* voting power of a voter under the existing Electoral College is:

The probability that the voter casts a decisive vote within his state
times
The probability that the state casts a decisive bloc of electoral votes in the Electoral College, *Given that* the voter is decisive within his state.

The probability that a voter casts a decisive vote in the state is essentially the probability that the state vote is tied, which is equal (to excellent approximation given a modestly large number n of voters) to $\sqrt{2/\pi n}$. The probability that the voter's state casts a decisive block of votes in the Electoral College is equal to the state's absolute Banzhaf power in the weighted voting game 51:538(270: 55, 34, ... , 3), i.e., one with 51 voters, a total weight of 538, a winning quota of 270, a weight of 55 for the largest player (California), 34 for the next largest (Texas), through 3 for the smallest state (Wyoming). The Banzhaf value for each state can be calculated using the appropriate algorithm (namely, *ipgenf*) from the *Computer Algorithms for Voting Power Analysis* website. Since (absolute) Banzhaf values are equivalent to the relevant probabilities, overall two-tier voting power for any voter is the product of these two quantities. Moreover, the probability that a state casts a decisive bloc of votes in the Electoral College is not conditional on the popular vote outcome within the state, so the condition '*given that* the voter is decisive within his state' in the formulation above is unnecessary.

In this manner, I could readily produce a chart such as Fig. 1 for my encyclopedia entry; it shows how individual voting power varies across states with different populations (based on the 2000 census). Since probabilities of individual decisiveness are very small, it is convenient to rescale voting power so that individual voting power in the least favored state (namely, Montana, the largest state with a single House seat) is set at 1.0 and in other states as multiples of this. The figure shows that voters in California have about 3.5 times the voting power of those in Montana. The two horizontal lines show mean individual voting power under the Electoral College and individual voting power under direct popular vote—the latter of course being the same for all voters and, perhaps surprisingly, substantially greater that mean voting power under the Electoral College (indeed, greater than the power of voters in every state other than most favored California).

Having completed my encyclopedia entry, I thought it would be interesting and straightforward to make similar charts for other variants of the Electoral College. The variants I considered fell into three categories: those that keep the state-level winner-take-all practice but use a different formula for apportioning electoral votes among states (e.g., basing electoral votes on House seats only, giving all states equal electoral votes, etc.), those that keep the existing apportionment of electoral votes but use something other than winner-take-all for the casting of state electoral votes, and a range of 'national bonus' plans.

All variants in the first category and also the *Pure District Plan* (under which each state is divided into as many equally populated electoral districts as it has electoral votes, and a candidate wins one electoral vote for each district carried) in the second category are simple two-tier systems, in which voting power calculations can be made in just the same way as for the existing Electoral College. The *Pure Proportional Plan* (under which each state's electoral votes are fractionally apportioned among candidates in a way that is precisely proportional to their popular vote shares in that state) and the *Whole Number Proportional Plan* (under which each state's electoral votes state are apportioned among the candidates on the basis of their popular vote shares, but in whole numbers using an apportionment formula in the manner of proportional representation electoral systems) require somewhat different but still straightforward calculations.[2] However, the *Modified District Plan* (under which a candidate wins one electoral vote for each Congressional District he carries *and* two electoral votes for each state he carries) and any *National Bonus Plan* (under which electoral votes are apportioned and cast as under the existing

[2]Such calculations reveal that the Modified District and Pure Proportional plans both give a substantial advantage to voters in small states (due to their advantage in the apportionment of electoral votes). The voting power implications of the Whole Number Proportional Plan are truly bizarre: voters in states with an even number of electoral votes have (essentially) zero voting power, while voters in states with an odd number of electoral votes have voting power as if electoral votes were equally apportioned among these states (Beisbart and Bovens 2008). All these findings are presented in Miller (2009) using charts similar to Fig. 1.

system but the candidate who wins the most popular votes nationwide is awarded a bonus of some fixed number additional electoral votes) present special difficulties. This is because each voter casts a single vote that counts in two ways: in the voter's district and state under the Modified District Plan, and in the voter's state and the nation as a whole under the National Bonus Plan. This means that the probability that a state casts a decisive pair of votes in the Electoral College (under the Modified District Plan), or the bonus is decisive (under the National Bonus Plan) depends in some degree on whether the voter casts a decisive vote at the district or state level respectively. In this event, the condition '*given that* the voter is decisive within his state' in the earlier formulation of double decisiveness is now necessary (at least in principle—one might speculate that it would make little difference in practice).

In his original work on voting power in the Electoral College, Banzhaf (1968) attempted to calculate individual two-tier voting power under the Modified District Plan by (1) calculating individual voting power through the voter's district, (2) separately calculating individual voting power through the voter's state, and then (3) adding these two probabilities together. Figure 2 displays voting power under the Modified District Plan (based on the 2000 census) when calculated in the Banzhaf manner. While the *relative* voting power of voters in different states appears reasonable and turns out to be approximately correct, Fig. 2 displays a major anomaly in that mean individual voting power exceeds individual voting power under direct popular vote. This is anomalous because Felsenthal and Machover (1998, pp. 58–59) demonstrate that, within the class of ordinary voting systems, mean individual voting power under direct popular vote maximizes the total Banzhaf score of all voters and therefore also maximizes mean voting power. This anomaly was not evident in Banzhaf's work, because he reported only relative voting power across states and never made comparisons of absolute individual voting power across Electoral College variants or with the direct popular vote system.[3]

More recent work by Edelman (2004) clarifies the nature of this problem but does not itself point to a solution. Edelman argued that individual voting power in two-tier voting systems of a representative nature (e.g., council or legislature) can be enhanced by providing some at-large representation in addition to single-member district representation. Edelman further showed that if voters cast separate and independent votes for their district and at-large representatives, and if the at-large representatives are elected on winner-take-all slates and vote as a bloc in the top tier, individual voting power may be determined by separately calculating individual voting power through the voter's district representation and through at-large representation and the adding the two probabilities together (essentially as Banzhaf tried to calculate voting power under the Modified District Plan). Edelman further shows that individual voting power so calculated is maximized when the number of at-large representatives is equal to the (approximate) square root of the

[3]Recalculation of Banzhaf's (1968) results shows that the same anomaly existed under the 1960 apportionment of electoral votes.

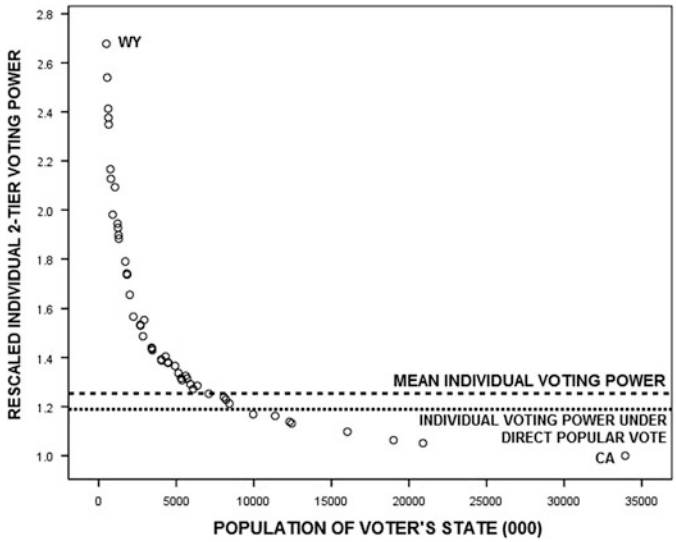


Fig. 2 Individual voting power by state population under the Modified District Plan (Banzhaf calculations)

total number of representatives and that such voting power exceeds individual voting power when all members are elected at-large.

The key assumption in Edelman's analysis is that voters cast *separate and independent votes* for district and at-large representation. Edelman claims that allowing separate and independent votes gives a voter more power because "he has more flexibility in the way he casts his vote." In many contexts, greater "flexibility" in casting votes may be valuable to voters, but only if possible election outcomes have multiple attributes that voters care about, e.g., if voters care, not only about what party controls the council, but also about its ideological balance, ethnic diversity, geographical representation, etc. But the foundational assumption of standard voting power theory is that "the measurement of voting power … concerns any collective body that makes *yes-or-no decisions* by vote" (Felsenthal and Machover 1998, p. 1; emphasis added), i.e., the setup is based on votes and outcomes that are both binary in nature. Edelman himself notes that the assumption of separate and independent votes does not apply to the Modified District Plan for Electoral College in which a voter casts a single vote that counts in two ways, though he speculates that, if the number of voters is large enough, voting power under this plan may be just about the same as when votes are separate and independent. In any event, even if the Modified District Plan or the National Bonus were modified to *allow* separate and independent votes, voters would never have reason to *use* their new-found "flexibility" to "split" these votes, given the binary nature of Presidential election outcomes—that is to say, there is no reason to vote for a Democratic-pledged elector at the district (or state) level and a Republican pledged-elector at the state (or national) level (or vice versa).

This gives us some insight into why the Banzhaf-style calculations for the Modified District and National Bonus Plans allows mean individual voting power to exceed what it would be under direct popular vote—they in effect assume, not only that voters *can* "split" their district (or state) and state (or national) votes in this manner, but also that they *actually do* "split" their votes *half the time*, thereby removing the correlation that would otherwise exist between district (or state) and state (or national) votes.

2 A Simple Example

As a warm-up exercise, let us consider the simplest case in which nine voters are partitioned into three uniform districts. Elections are held under four distinct voting rules, each of which is symmetric with respect to both voters and two candidates A and B. Under all rules, voters cast a single vote that counts in two ways, i.e., first in the 'district' part of the upper-tier and second in the 'at-large' part of the upper-tier. With the U.S. Electoral College in mind, we may refer to lower-tier votes as 'popular votes' and upper-tier votes as 'electoral votes.' These are the four voting rules:

1. *Pure District System*: each district casts one electoral vote, and the candidate winning a majority of electoral votes (two out of three) is elected;
2. *Small At-Large Bonus System*: each district casts one electoral and one additional electoral vote is cast at-large, and the candidate winning a majority electoral votes (three out of four) is elected (ties may occur in the upper tier);
3. *Large At-Large Bonus System*: each district casts one electoral vote and a 'winner-take-all' bloc of two electoral votes is cast at-large, and the candidate winning a majority of electoral votes (three out five) is elected; and
4. *Pure At-Large System*: there no districts or, in any case, a bloc of 4 or more electoral votes is cast at-large, so the districts are superfluous and the candidate winning a majority of the popular votes (five out of nine) is elected.

Let us consider things from the point of view of a focal voter i in District 1, who confronts $2^8 = 256$ distinct combinations of votes that may be generated by the other eight voters. We want to determine, for each voting rule, in how many of the 256 combinations voter i is *decisive*, in the sense that i's vote tips the election outcome one way or the other.[4] The number of such combinations is voter i's

[4] If the number of voters n is even (e.g., $n = 100$), the *interpretation* of a decisive vote differs somewhat according to whether the voting context is parliamentary or electoral. Under usual parliamentary rules, a tie vote defeats a motion, so voter i is decisive in any voting combination in which 50 other voters vote 'yes' and 49 vote 'no,' as the motion passes or fails depending on whether i votes 'yes' or 'no.' However, in elections between two candidates (our present concern), voting rules are typically neutral between the candidates, so a tie outcome might be decided by the flip of a coin. In this event, a voter i is "half decisive" in any voting combination in which 50 other voters vote for A and 49 for B (A wins if i votes for A and each candidate wins with 0.5 probability

Banzhaf score, and the number of such combinations divided by 256 is voter *i*'s (absolute) *Banzhaf voting power* in the two-tier voting game. If each combination is equally likely, voter *i*'s Banzhaf power is equal to the *probability* that *i* casts a vote that is *doubly decisive*, which can occur in three ways: (1) the individual vote is decisive in *i*'s district *and* the district vote is decisive in the upper tier, (2) the individual vote is decisive in the at-large component *and* the at-large bloc is decisive in the upper tier, and (3) the individual vote is decisive in both *i*'s district and the at-large component *and* these combined votes are decisive in the upper tier.

Table 1 accounts for all 256 possible voting combinations by listing and enumerating each of the 58 distinct (anonymous) vote profiles giving rise to each combination and indicates for each whether voter *i*'s vote is decisive under each of the four rules. At the bottom, Table 1 reports voter *i*'s Banzhaf score and voting power for each rule. We see that Banzhaf voting power increases as the weight of the at-large component increases.[5] The bottom of the table shows Banzhaf voting power calculated (in the manner of Edelman) on the assumption that voters cast *separate and independent votes* at the district and at-large levels. In the Edelman setup, individual voting power is maximized with a mixture of district and at-large electoral votes such that the at-large component is approximately the square root of the total number of electoral votes. The Edelman setup does not generate an ordinary simple voting game, and therefore Edelman voting power values cannot be calculated in the manner of Table 1; however, they can be readily calculated, as shown in the third note at the foot of the table. Any district vote profile may occur in conjunction with any popular vote split and, in particular, a candidate can win the at-large vote without carrying any district.

Table 2 is derived from Table 1 and has two types of entries in each cell. First, it crosstabulates the 256 voting combinations with respect to whether voter *i*'s district vote (DV) is tied, thereby making *i*'s vote decisive within the district (column variable), and whether the popular (at-large) vote (PV) is tied, thereby making *i*'s vote decisive with respect to the at-large vote (row variable). We call each cell a *contingency*, and the lower number in each cell indicates number of voting combinations giving rise to that contingency. The contingencies themselves pertain to characteristics of the first-tier vote only. However, the four top numbers in each cell pertain to the four distinct upper-tier voting rules and indicate, for each voting rule, the number of combinations in which *i*'s vote is doubly decisive and that thereby contribute to *i*'s Banzhaf score.

---

if *i* votes for B) *and also* in any voting combination in which 49 other voters vote for A and 50 for B. The upshot is that voter *i*'s total Banzhaf score (and voting power) is the same under either interpretation. Thus we can (and will) speak loosely "the probability of a tie vote" even when the number of other voters is even. More obviously, we can (and will) speak interchangeably between "the probability of voter *i* breaking what would otherwise be a tie vote" and "the probability of a tie vote" when the number of voters is large.

[5]However, with only nine voters, the Large At-Large System with a single vote that counts in two ways is effectively equivalent to the Pure At-Large System, because the candidate who wins the at-large vote must win at least one district and thus three out of five electoral votes.

Table 1 All possible vote profiles confronting focal voter *i* in district 1, given nine voters uniformly partitioned into three districts

			Number of times voter *i* is decisive (total is *i*'s Bz score)			
Pop. vote	District vote profile	***k***[a]	(1) Pure district	(2) Small AL	(3) Large AL	(4) All AL
8–0	(2–0) (3–0) (3–0)	1	0	0	0	0
	Total	**1**	**0**	**0**	**0**	**0**
7–1	(1–1) (3–0) (3–0)	2	0	0	0	0
	(2–0) (2–1) (3–0)	3	0	0	0	0
	(2–0) (3–0) (2–1)	3	0	0	0	0
	Total	**8**	**0**	**0**	**0**	**0**
6–2	(0–2) (3–0) (3–0)	1	0	0	0	0
	(1–1) (2–1) (3–0)	6	0	0	0	0
	(1–1) (3–0) (2–1)	6	0	0	0	0
	(2–0) (3–0) (1–2)	3	0	0	0	0
	(2–0) (2–1) (2–1)	9	0	0	0	0
	(2–0) (1–2) (3–0)	3	0	0	0	0
	Total	**28**	**0**	**0**	**0**	**0**
5–3	(0–2) (3–0) (2–1)	3	0	0	0	0
	(0–2) (2–1) (3–0)	3	0	0	0	0
	(1–1) (3–0) (1–2)	6	6	3[b]	0	0
	(1–1) (2–1) (2–1)	18	0	0	0	0
	(1–1) (1–2) (3–0)	6	6	3[b]	0	0
	(2–0) (3–0) (0–3)	1	0	0	0	0
	(2–0) (2–1) (1–2)	9	0	0	0	0
	(2–0) (1–2) (2–1)	9	0	0	0	0
	(2–0) (0–3) (3–0)	1	0	0	0	0
	Total	**56**	**0**	**6**[b]	**0**	**0**
4–4	(0–2) (3–0) (1–2)	3	0	1.5[b]	3	3
	(0–2) (2–1) (1–2)	9	0	4.5[b]	9	9
	(0–2) (1–2) (3–0)	3	0	1.5[b]	3	3
	(1–1) (3–0) (0–3)	2	2	2	2	2
	(1–1) (2–1) (1–2)	18	18	18	18	18
	(1–1) (1–2) (2–1)	18	18	18	18	18
	(1–1) (0–3) (3–0)	2	2	2	2	2
	(2–0) (2–1) (0–3)	3	0	1.5[b]	3	3
	(2–0) (1–2) (1–2)	9	0	4.5[b]	9	9
	(2–0) (0–3) (2–1)	3	0	1.5[a]	3	3
	Total	**70**	**40**	**55**	**70**	**70**
3–5	**Dual of 5–3**	**56**	**12**	**6**	**0**	**0**
2–6	**Dual of 6–2**	**28**	**0**	**0**	**0**	**0**
1–7	**Dual of 7–1**	**8**	**0**	**0**	**0**	**0**
0–8	**Dual of 8–0**	**1**	**0**	**0**	**0**	**0**
	Total [Bz score]	**256**	**64**	**67**	**70**	**70**
	Bz power		**0.25**	**0.26172**	**0.27344**	**0.27344**
	Edelman Bz power[c]		**0.25**	**0.29004**	**0.33008**	**0.27344**

[a]***k*** is the number of distinct voter combinations giving rise to the specified district vote profile

[b]In these profiles, Banzhaf awards voter *i* "half credit," as *i*'s vote is decisive with respect to whether a particular candidate wins or there is a tie between the two candidates. (Under the other voting rules, ties cannot occur.)

[c]Edelman Bz power = Prob. decisive in district × Prob. district decisive in Tier 2 + Prob. decisive at-large × Prob. at-large decisive in Tier 2

AL = 1: $5 \times 0.375 + 0.27344 \times 0.375 = 0.29004$

AL = 2: $0.5 \times 0.25 + 0.27344 \times 0.7 = 0.33008$

Table 2 Summary of Table 1 identifying Contingencies 1–3

	DV tied	DV not tied	Total
PV tied	40/40/40/40	0/15/30/30	40/55/70/70
	Contingency 1	***Contingency 2***	
	40	30	70
PV not tied	24/12/0/0	0/0/0/0	24/12/0/0
	Contingency 3		
	88	98	186
Total	64/52/40/40	0/15/30/30	64/67/70/70
	128	128	256

Pure district/1 A-L/2 A-L/All A-L

The numbers in Table 2 were determined by consulting Table 1, and Table 1 in turn was easy (if tedious) to construct. But if the number of voters expands even slightly, it becomes impractical to replicate Table 1 (for example, with 25 voters the number of possible combinations facing voter i is $2^{24} = 16{,}777{,}216$), so some less direct method for enumerating (or estimating) Banzhaf scores and voting power values must be devised. We now turn to a larger-scale example, though still simplified relative to either Electoral College variant.

3 A Large-Scale Example with Uniform Districts

We now consider an example in which $n = 100{,}035$ voters are uniformly partitioned into $k = 45$ districts with 2,223 voters, each with a single electoral vote and with a bloc of 6 additional electoral votes elected at-large.

We note two relevant baselines. Given 51 districts and no at-large seats and using the standard approximation $\sqrt{2/\pi n}$, with $n = 100{,}035/51 = 1{,}961.47$, for the probability of a tie vote, individual voting power within a district is 0.0180156. Using the Leech website, the voting power of each district in the second tier is 0.112275. Thus individual voting power (the probability of double decisiveness) is $0.0180156 \times 0.112275 = 0.0020227$. At the other extreme, with 25 or fewer districts (i.e., effectively direct popular vote), individual voting power is simply $\sqrt{2/\pi n}$, with $n = 100{,}035$, or 0.0025227.

We begin with Table 3, set up in the same manner as Table 2 and initially pertaining to lower-tier votes only. Since the number of voting combinations is impossibly large, *proportions* rather than counts of combinations are displayed and, given random voting, these are also *probabilities*. We first calculate the probability that the popular vote is tied, which gives us the total in the first row. As noted just above, this probability is 0.0025227. Using the same approximation with $n = 100{,}035/45 = 2{,}223$, we calculate the probability that the vote in i's district is tied to be 0.0169227, which gives us the first column total. Subtraction from 1.0000000 gives us the totals in the second row and second column.

Table 3 Marginal proportions in large-scale example

	DV tied	DV not tied	Total
PV tied	***Contingency 1***	***Contingency 2***	0.0025227
PV not tied	***Contingency 3***		0.9974773
Total	0.0169227	0.9830773	1.0000000

Table 4 Contingency proportions in large-scale example plus Edelman calculations

	DV tied	DV not tied	Total
PV tied	0.0000426	0.0024900	$\times 0.628702 =$ **0.001586** 0.0025227
PV not tied	0.0168801	0.9805972	0.9974773
Total	$\times 0.080083 =$ **0.0013552** 0.0169227	0.9830773	**0.0029412** 1.0000000

So far as Edelman-style calculations are concerned, we are almost done. If district and at-large votes are separate and independent, we can calculate the probabilities of contingencies simply by multiplying the corresponding row and column probabilities, as shown in Table 4. But, given Edelman's assumptions, we need not be concerned with the interior cells at all. We need look only at the marginal proportions in the first row and first column and then take account of voting in the upper tier. Upper-tier voting is given by the voting rule 46:51(26:6,1, ... ,1)—that is, it is a weighted voting game with 46 players (45 districts plus the at-large bloc), a total of 51 electoral votes, a quota of 26 (a bare majority of the total of 51 electoral votes), and voting weights of 6 for the at-large bloc and 1 for each district. The Leech website produces 0.628702 and 0.080083 as the voting power for the at-large bloc and each district respectively. The voting power of voter i through district representation is his probability of being decisive within his district times the probability that is district is decisive in the second tier, i.e., $0.0169227 \times 0.080083 = 0.0013552$, and i's voting power through at-large representation is his probability of being decisive in the popular vote times the probability that the at-large bloc is district is decisive in the second tier, i.e., $0.0025227 \times 0.628702 = 0.0015860$. Within Edelman's setup, the overall voting power of each voter is simply the sum of these probabilities, i.e., 0.0029412, as also shown in Table 4.[6] Note that this is greater than voting power under direct popular election, i.e., 0.0025227. Figure 3 shows Edelman-style voting power for all magnitudes of at-large representation, illustrating Edelman's result that such voting power is maximized when the size of the at-large component is set at the square root of the total size of the representative body.

[6]Taking the sum of the voting powers associated with each of the voter's (district and at-large) votes may appear to double-count those voting combinations in Contingency 1 in which *both* of i's two votes are doubly decisive, but at the same time it misses voting combinations in Contingency 1 in which *neither* vote by itself is doubly decisive but the two votes together are, and it turns out that these combinations exactly balance out (Beisbart 2007).

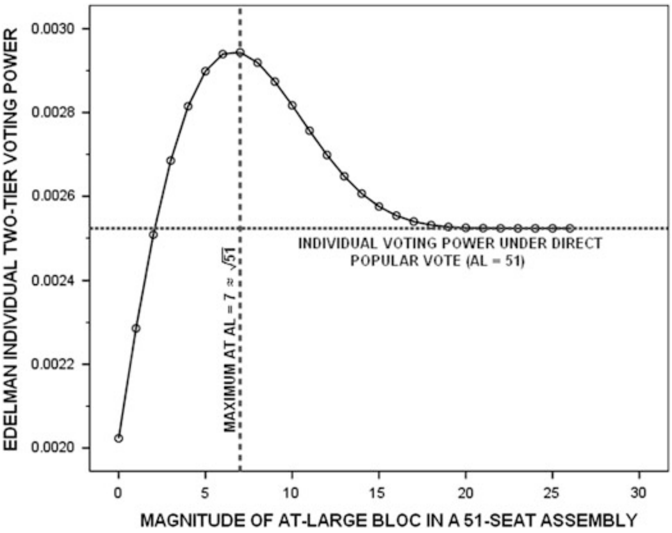


Fig. 3 Individual voting power by magnitude of the at-large bloc (Edelman calculations)

Table 5 Marginal and contingency probabilities with one vote counting in two ways

	DV tied	DV not tied	Total
PV tied	0.0025512 ⇓ 0.0000431	0.0024796	0.0025227
PV not tied	0.0168796	0.9805977	0.9974773
Total	0.0169227	0.9830773	1.0000000

If, in contrast to the Edelman setup, each voter has a single vote that counts for both district and at-large representation, we have an ordinary simple voting game, and individual two-tier voting power cannot exceed the 0.0025227 level resulting from direct popular (Pure At-Large) election (Felsenthal and Machover 1998, pp. 58–59). However, voting power calculations become far more complex.

We first return to Table 3 and observe that, in the single-vote setup, the marginal probabilities are the same, as is shown in Table 5. However, the fact that voters cast the same vote for both district and at-large representation induces a degree of correlation between the vote in any district and the at-large vote, so the probability that both votes are tied is greater than the 0.0000426 in the Edelman setup.

We can directly calculate the *conditional probability* that the at-large vote is tied given that a district vote is tied. Given that the vote in i's district is tied, the overall at-large vote is tied if and only if there is also a tie in the residual at-large vote after the votes cast in voter i's district are removed. The probability of this event is given by the standard approximation $\sqrt{2/\pi n}$, where n is now $100{,}035 - 2{,}223 = 97{,}812$, and is equal to 0.0025512 as shown in Table 5. We can now derive the unconditional probability that both types of ties occur simultaneously by multiplying this conditional probability by the probability that the district vote is tied in the first place, i.e., $0.0025512 \times 0.0169227 = 0.0000431$. With this piece of the puzzle in place, the probabilities of the other contingencies are determined by subtraction. Comparing Tables 5 and 4, we observe that the probabilities of the contingencies differ only slightly, with the probability of ties at one level but not the other being slightly less in the single-vote setup, so the substantially lower overall voting power arising from this setup relative to Edelman's evidently results mostly from the workings of upper-tier voting.

In any event, voter i is decisive in the two-tier voting process only if the at-large and district votes are both tied (Contingency 1), the at-large vote only is tied (Contingency 2), or the district vote only is tied (Contingency 3). Having determined the probabilities of these contingencies, our next—and much more difficult—task is to determine, given each of these contingencies, the probability that voter i's vote is decisive in the upper tier as well.

First, let's form some general expectations. Contingency 1, being the conjunction of two already unlikely circumstances, is extraordinarily unlikely to occur but, if it does occur, voter i is very likely to be doubly decisive. Voter i is doubly decisive if and only if neither candidate has won a majority of 26 electoral votes from the 44 other districts—put otherwise, if each candidate has won between 19 and 25 districts. By breaking a tie in both his district and at-large vote, voter i is tipping 7 electoral votes one way or the other, thereby giving one or other candidate the 26 electoral votes required for election. Given random individual voting, the electoral votes of the other 44 districts are likely to be quite evenly divided. Since each candidate is likely to have won about half of them, it likely that neither has won as many as 26 out of 44 districts, thereby making voter i doubly decisive.

Contingency 2 is considerably more likely to occur than Contingency 1, while voter i's probability of double decisiveness is only slightly less, since the voter is tipping almost as many electoral votes (six rather than seven) one way or the other in Contingency 2 as in Contingency 1. Voter i is now doubly decisive if and only if neither candidate has won a majority of 26 electoral votes from all 45 districts—put otherwise, if each candidate has won between 20 and 25 districts. By breaking an at-large vote tie, voter i is tipping 6 electoral votes one way or the other and thereby gives one or other candidate the 26 electoral votes required for election. Again, given random voting, the electoral votes of the 45 districts are likely to be quite evenly divided, so it quite likely that neither candidate has won as many as 26 districts.

Contingency 3 is still more likely to occur than Contingency 2, but voter i is far less likely to be doubly decisive in this contingency, since he is tipping only a single electoral vote one way or the other. Voter i is doubly decisive if and only if neither candidate has won a majority of 26 electoral votes from the other 44 districts and the at-large bloc of 6 votes, i.e., in the event that there is an overall 25–25 electoral vote tie. Such a tie results if and only if one candidate has carried 25 districts, while the other candidate has carried 19 districts *and* the at-large vote. The probability of such an event is very small for three reasons:

1. An *exact* tie in the second-tier electoral vote tie is required, because i is tipping only a single electoral vote;
2. The split in *district* electoral votes must be unequal in a degree that depends on the number of at-large seats (here 25–19 with 6 at-large seats) in order to create a tie in overall electoral votes, and such an unequal split is less likely than an equal split, since random voting always produces 50–50 expectations; and
3. This rather unlikely 25–19 split in favor of one candidate in terms of district electoral votes must come about in the face of a popular vote majority in favor of the *other* candidate (which earned him the at-large bloc).

The last point implies that, in Contingency 3, voter i is doubly decisive only if i's vote can bring about the kind of *election inversion* (or 'reversal of winners,' 'wrong winner,' 'referendum paradox,' etc.) in which the candidate who wins with respect to district electoral votes at the same time loses with respect to overall at-large (popular) votes (Miller 2012). It is characteristic of districted election systems such as U.S. Presidential elections and U.K. general elections that such election inversions may occur, but they are quite unlikely unless the (at-large or popular vote) election is very close. But we must bear in mind that almost all large-scale random elections are extremely close. Indeed, if district and at-large votes are cast separately and independently in the Edelman manner so there is no correlation between them, it is evident that 50 % of all random elections produce election reversals. This is shown in Fig. 4a, which is based on a sample of 30,000 random elections in which the at-large vote and the district votes were generated independently. In contrast, when the popular vote is the district vote summed over all districts, a substantial correlation is induced between district and at-large votes, which considerably reduces the incidence of election inversions. This is shown in Fig. 4b, which is based on the same sample of 30,000 random elections when the at-large popular vote is the sum of the district votes. In this sample, election inversions occurred in 20.4 % of the elections, very closely matching the rate of 20.5 % found by Feix et al. (2004) in a sample of one million random (or 'Impartial Culture') elections.[7]

4 Random Election Simulations

Having formed expectations about the probability of double decisiveness in each contingency, we must now assign numbers to these probabilities. While it may be possible to proceed analytically, I have found the obstacles to be formidable and have instead proceeded on the basis of large-scale simulations. For the present

[7]The correlation between the number of uniform districts carried by a candidate and the candidate's national popular vote is about +0.784. This degree of associations appears to be essentially constant regardless of the number of voters or districts, provided the latter is more than about 20 and the former is more than a thousand or so per district.

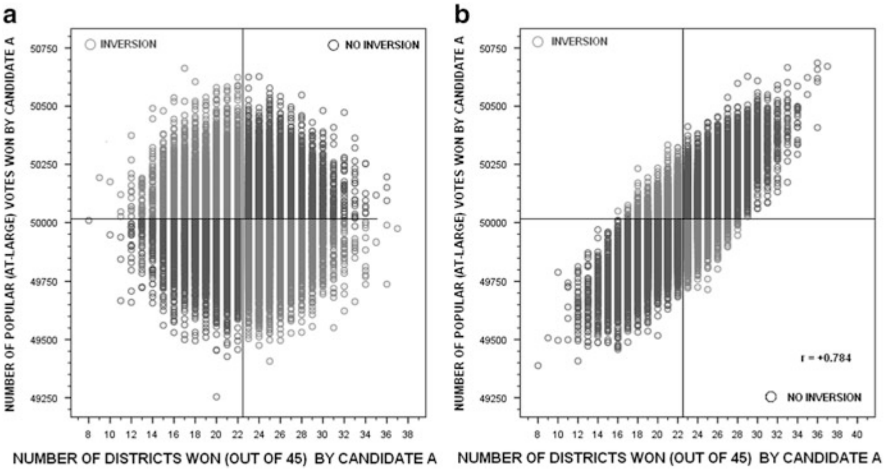


Fig. 4 Two-tier random election outcomes. (**a**) Separate and independent votes in each tier (Edeleman). (**b**) When one vote counts the same way in both tiers

case with 45 districts and 6 at-large seats, I have generated a sample of 1.2 million random elections.[8]

The next question is how to use the results of the simulation to estimate the relevant probabilities. The most direct approach is to produce the crosstabulation depicted in Table 6, which shows the absolute frequencies produced by these simulations. The number in the lower part of each cell is the number of times that the contingency arose. The number in the upper part of each cell is the number of times voter i was doubly decisive in that contingency. Overall, voter i was doubly decisive (DD) in 2,970 elections out of 1,200,000. Thus the estimated *a priori* voting power of voter i (and every other voter, given the overall symmetry) is 2,970/1,200,000 or 0.002475, a figure that sits comfortably between the lower bound of 0.0020227 for district only voting and the upper bound of 0.0025227 for direct popular voting. Our confidence in the simulated elections is reinforced by comparing Table 7, in which all absolute frequencies in Table 6 are converted into proportions (and estimated probabilities), with Table 5. It is evident that the relative frequency of each contingency closely matches the exact probabilities calculated earlier.

A second approach is to replace the estimated probabilities of each contingency in the lower part of each cell in Table 7 by the known probabilities displayed

[8] As before, each of the 45 districts has 2,223 voters, a number selected so that both district and at-large vote ties may occur before focal voter i (in District 1) casts his vote and so that no ties occur after i has voted. The simulations, which are generated by SPSS syntax files, operate at the level of the district: the vote for candidate A in each district is a number drawn randomly from a normal distribution with a mean of $2{,}223/2 = 1{,}111.5$ and a standard deviation of $\sqrt{.25 \times 2223}$, i.e., the normal approximation to the binomial distribution with $p = 0.5$, and then rounded to the nearest integer.

Table 6 Crosstabulation of district and at-large ties in 1.2 million random elections (case counts)

	DV tied	DV not tied	Total
PV tied	49	2,554	2,603
	***Prob. of DD* = 0.960784**	***Prob. of DD* = 0.862838**	
	51	2,960	3,011
PV Not tied	367	0	367
	***Prob. of DD* = 0.018198**		
	20,167	1,176,822	1,196,989
Total	416	2,554	2,970
			***Prob. of DD* = 0.002475**
	20,218	1,179,782	1,200,000

Table 7 Crosstabulation of district and at-large ties in 1.2 million random elections (proportions)

	DV tied	DV not tied	Total
PV tied	0.0000408	0.0021283	0.0021692
	0.0000425	0.0024667	0.0025092
PV Not tied	0.0003058	0.0000000	0.0003058
	0.0168058	0.9806850	0.9974908
Total	0.0003467	0.0021283	**0.0024750**
	0.0168483	0.9831517	1.0000000

in Table 5. In this case, the numbers are so similar that the substitution makes essentially no difference, as voter *i*'s estimated voting power becomes 0.0024880, in contrast to 0.0024750 using simulated data only.

A third approach is suggested if we examine the frequency distributions underlying the cells in Table 6. With respect to Contingency 1, Fig. 5a shows the frequency distribution of districts won by Candidate A in the 51 elections in which both the at-large vote and the vote in an (arbitrarily selected) District 1 are tied. A voter in District 1 is doubly decisive provided that the number of districts won by either candidate lies within the range of 19–25. This was true in 49 elections out of the 51 elections, giving voter *i* a 0.960784 probability of double decisiveness in this contingency. But it evident that another sample of 1.2 million random elections (including about 50 belonging to Contingency 1) might produce a substantially different statistic. And, given a larger sample size, we would expect this distribution to fit a more or less normal pattern, rather than the bimodal pattern that happens to appear in Fig. 5a. So, given the present sample of elections, a more reliable estimate of voter *i*'s probability of double decisiveness may be derived by supposing that the underlying distribution of districts won by Candidate A is normally distributed with a known mean of 22 (i.e., one half of the other 44 districts), rather than the sample statistic of 22.294118, and with the standard deviation of 2.032674 found in this sample. (From this point of view, the main purpose of the simulation is to get an estimate of this standard deviation.) The estimated proportion of times voter *i* is doubly decisive is therefore equal the proportion of the area under a normal curve that lies within 3.5/2.032674 = 1.72187 standard deviations from the mean, which is 0.914907. This suggests that the direct result of the simulation of 0.960785 is too

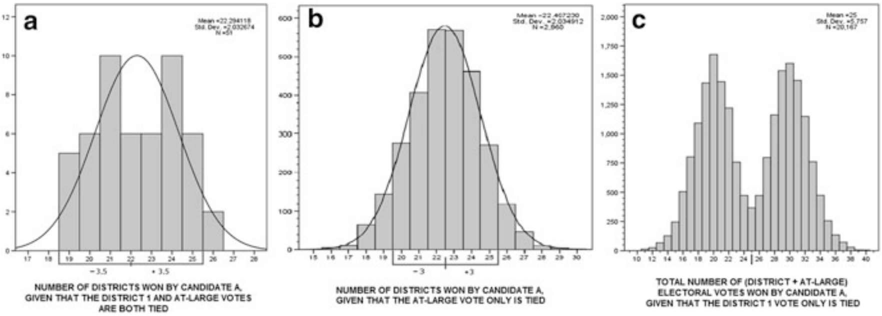


Fig. 5 Distribution of electoral votes won by Candidate A. (**a**) In Contingency 1. (**b**) In Contingency 2. (**c**) In Contingency 3

high, and indeed Fig. 5a suggests that it was only by 'good luck' that Candidate A never won fewer than 19 districts.

A fourth approach—which is especially appealing with respect to Contingency 1—is to exploit the symmetry resulting from the uniformity of districts and to use the simulated data set to make parallel calculations for voters in all 45 districts and average them. This results in a rate of double decisiveness of 0.919027, slightly higher than the previous result, and this is probably the best estimate given the simulated data set.

In like manner, Fig. 5b shows the frequency distribution of districts won by Candidate A in the contingency that the at-large vote only is tied. The actual distribution closely matches a normal distribution with a mean of 22.5 (i.e., one half of all the 45 districts). Given the much larger (2,960) sample of elections in Contingency 2, it is unsurprising that the normal curve approach to estimating voter *i*'s double decisiveness produces essentially the same result (0.859592) as the sample statistic itself (0.862838), and in this case the statistic is probably more reliable.

Figure 5c shows the frequency distribution of (district plus at-large) electoral votes won by Candidate A in the contingency that the vote in District 1 only is tied. Since the distribution is clearly bimodal (resulting from the fact that the at-large vote is not tied, as in Fig. 5a, b, and one or other candidate has won the block of six at-large votes), we obviously cannot use the normal curve approach. However, Contingency 3 is by far the most likely of the three contingencies that allow voter *i* to be doubly decisive, so the sample size is very large ($n = 20{,}167$) and the sample statistic for a 25–25 electoral vote tie ($367/20{,}167 = 0.018198$) should be highly reliable.

Putting this altogether in Table 8, by pooling the results from all districts to estimate probability of double decisiveness in Contingency 1, using the sample statistics in Contingencies 2 and 3, and using the known probabilities for the contingencies themselves, we get an estimate of voter i's voting power of 0.0024863, compared with 0.0024750 using sample statistics only (and 0.0024880 using the sample statistics for probabilities of decisiveness in conjunction with the known probabilities for the contingencies themselves). In sum, we can be pretty confident that the true value of voter *i*'s voting power is just about 0.0024786, putting it

Table 8 Final estimate of individual voting power in large scale example

	DV tied	DV not tied	Total
PV tied	×0.919027 = **0.0000397**	×0.862836 = **0.0021394**	0.0021791
	0.0000432	0.0024795	0.0025227
PV Not tied	×0.018198 = **0.0003072**	0.0000000	0.0003058
	0.0168795	0.9805978	0.9974773
Total	0.0003467	0.0021314	**0.0024863**
	0.0169227	0.9830773	1.0000000

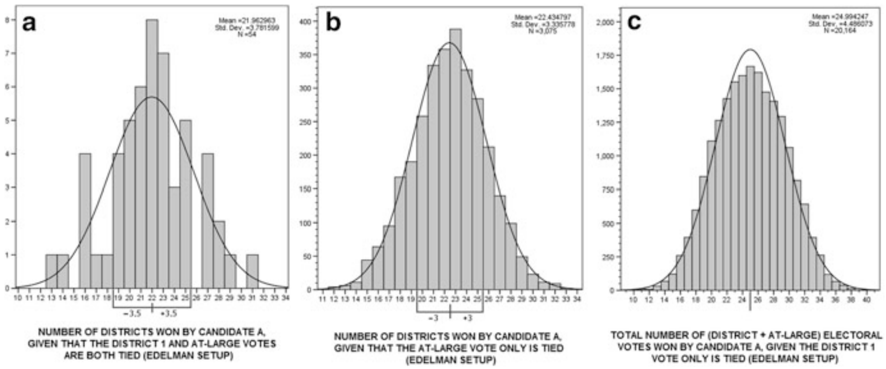


Fig. 6 Distribution of electoral votes won by candidate A (Edelman setup). (**a**) In Contingency 1. (**b**) In Contingency 2. (**c**) In Contingency 3

slightly but clearly below the value of 0.002523 that results from direct popular vote. This contrasts of the Edelman value of 0.0029412 that results when voters cast separate and independent votes at the district and at-large levels.

Comparing Fig. 5a–c with Fig. 6a–c that results in the Edelman setup makes evident how the Edelman setup produces a greater probability of double decisiveness. We see that each contingency occurs with virtually the same probability in the two setups (as we saw before in the calculations displayed in Tables 4 and 5). In the first two contingencies, a voter is actually less likely to be doubly decisive in the Edelman setup, as the spread in districts won by either candidate is substantially larger. This results from the absence of a correlation between popular votes won and number of districts won that results when each voter casts a single vote that counts twice (Fig. 4b) rather that two separate and independent votes (Fig. 4a). But this effect is more than wiped out in Contingency 3, where two setups result in quite different distributions of electoral votes won. In the single-vote setup, the distribution is strikingly bimodal (the distance between the modes depending on the number of at-large electoral votes relative to the total) because, as a candidate wins more districts, he is more likely to win the at-large vote as well, whereas in the Edelman setup no such correlation exists. Given the parameters we are working with (6 at-large electoral votes out of 51), the Edelman setup produces a distribution that is unimodal but, relative to a normal curve, slightly 'squashed' in the center

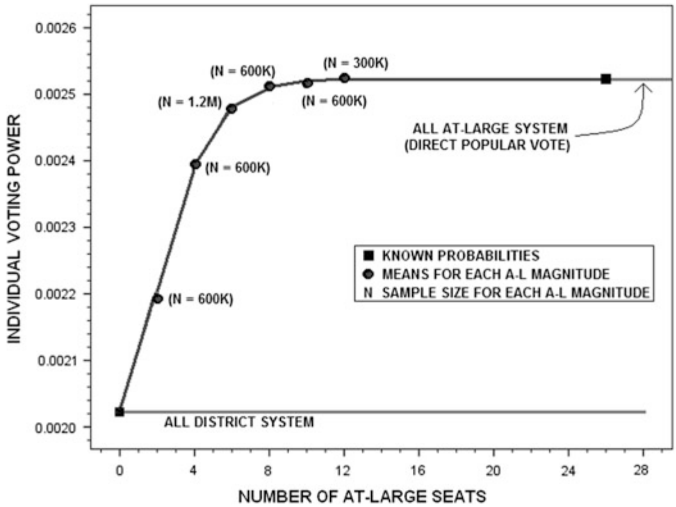


Fig. 7 Individual voting power by magnitude of at-large component

(Fig. 6c). If the relative magnitude of the at-large component were increased, the 'squashing' effect would be increased and would in due course produce bimodality, but it would always be substantially less than in the single-vote setup with the same at-large component. Thus, unless at-large component is wholly controlling (e.g., 26 electoral votes out of 51), the Edelman setup makes an even split of electoral votes far more likely than does the single-vote setup and thereby greatly enhances the probability of double decisiveness in Contingency 3, which in turn is by far the most probable contingency that (in either setup) allows double decisiveness.

I have duplicated the same kinds of simulations, with varying sample sizes, for other odd values of the at-large component within a fixed total of 51 electoral votes. The results (with sample sizes) are displayed in Fig. 7.[9] The general pattern of the relationship between the magnitude of the at-large component and individual power is very clear and is in sharp contrast with the pattern of the same relationship in the Edelman setup shown in Fig. 3.

5 The National Bonus Plan for the U.S. Electoral College

The previous analysis pertained to voting systems with uniform districts, all of which have the same number of voters and electoral votes. The most direct Electoral College application of the kind of analysis set out above pertains to variants of the

[9]The vertical axis in Figs. 7 and 9a, b must show actual, rather than rescaled, voting power, because the voting power of the least favored voter varies as the bonus (at-large) component varies.

National Bonus Plan, under which 538 electoral votes are cast in the present manner but the national popular vote winner is awarded a bonus of some number of ('at-large') electoral votes.[10] However, in this case the 'districts' (i.e., the states) are not uniform, having different numbers of both voters and electoral votes.

Like the previous example, under the National Bonus Plan votes count in two distinct upper tiers, i.e., the voter's state and the nation as a whole, with the result that doubly decisive votes can arise in three distinct contingencies: (1) a vote is decisive at both the state and national levels and the combination of the state's electoral votes and the national bonus is decisive in the Electoral College; (2) a vote is decisive at the national level only and the national bonus is decisive in the Electoral College; and (3) a vote is decisive at the state level only and the state's electoral votes are decisive in the Electoral College. However, under the bonus plan, the relevant probabilities and simulation estimates must be separately determined for voters in each state, each with its own number of voters and electoral votes. While the calculations and simulations are in this respect more burdensome, the procedure is a straightforward extension of that set out in the previous section. The following simulation results were based on a sample of 256,000 random elections, each with about 122 million voters.[11]

Figure 8a displays individual voting power, when calculated in the Banzhaf/Edelman manner, under a National Bonus Plan with a bonus of 101 electoral votes for the national popular vote winner. At first blush, Fig. 8a may look very similar to Fig. 1 for the existing Electoral College. But inspection of the vertical axis reveals that the inequalities between voters in large and small states are considerably compressed relative to the existing system. Moreover, the same anomaly occurs here as with Banzhaf's calculations for the Modified District Plan, in that mean individual voting power (considerably) exceeds that under direct popular vote. Figure 8b displays individual voting power with a 101 electoral vote national bonus calculated in the manner set out in Sect. 4.[12]

Figure 9a displays individual voting power with a national bonus of varying magnitude, again calculated in the Banzhaf/Edelman manner, while Fig. 9b shows

[10]Though this idea had been around earlier, it was most notably proposed by Schlesinger (2000) following the 2000 election. He proposed a national bonus of 102 electoral votes—two for each state plus the District of Columbia. However, given an even number (538) of 'regular' electoral votes, it would seem sensible to make the bonus an odd number in order to definitively eliminate the possibility of electoral vote ties (though ties would be far less likely than at present given any substantial nation bonus). It is clear that the motivation for a national bonus is to reduce the probability of election inversions, not to redistribute voting power.

[11]Given such a large electorate size, few if any elections were tied at the state or national level, so electoral vote distributions were taken from a somewhat wider band of elections, namely those that fell within 0.2 standard deviations of an precise tie in the state or national popular vote. (Random elections with many voters are very close, so the standard deviation is very small. Moreover, the ordinate of a normal curve at a standard score of ± 0.2 is about 0.98 times that at a standard score of zero, so the density of elections is essentially constant in the neighborhood of a tie.)

[12]The plotted points in Fig. 8b, unlike those in Fig. 8a, are estimates subject to some sampling error, but its effects are probably invisible.

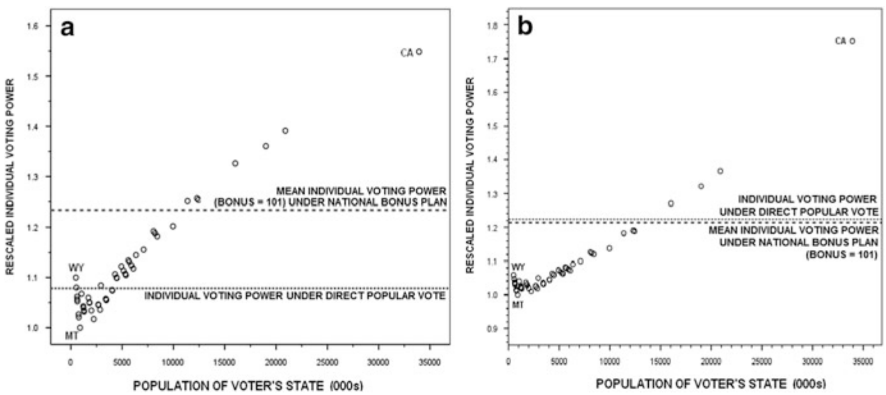


Fig. 8 Individual voting power by state population under the National Bonus Plan (Bonus = 101). (**a**) By Banzhaf–Edelman calculations. (**b**) By present calculations

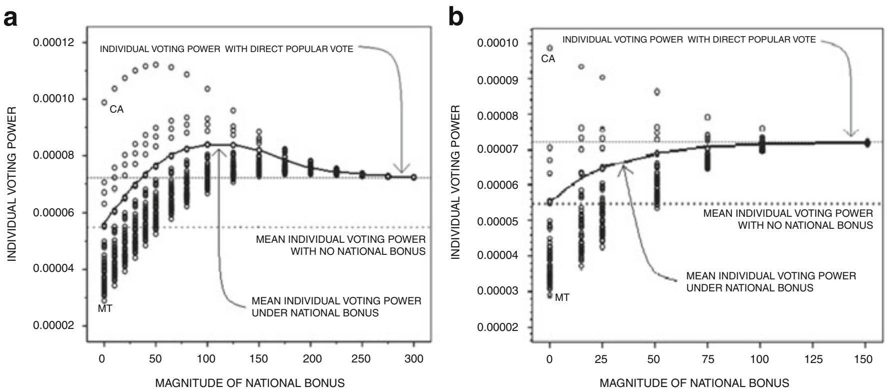


Fig. 9 Individual voting power by state population by magnitude of national bonus. (**a**) By Banzhaf–Edelman calculations. (**b**) By present calculations

the same when voting power is measured in the manner set out in Sect. 4. A bonus of zero is equivalent to the existing Electoral College system and a bonus of at least 533 (like an at-large component of four or more electoral votes in the simple example considered in Sect. 3) is logically equivalent to direct popular vote.[13] However, Fig. 9b indicates that a bonus greater than about 150 is essentially equivalent to direct popular vote.

A comparison of Figs. 8a and 8b indicates that, under the National Bonus Plan with a bonus of 101 electoral votes, the relative voting power of voters in

[13]With each vote counting the same way at the state and national levels, the national popular vote winner must win at least one state with at least three electoral votes, and 533 is the smallest number B such $B + 3 > 0.5(538 + B)$.

different states as calculated here is about the same—though small-state voters are slightly more favored—as under the Banzhaf–Edelman calculations, but the latter considerably overestimate voters' absolute voting power.

6 The Modified District Plan for the U.S. Electoral College

Under the Modified District Plan, a candidate wins one electoral vote for each Congressional District he carries and two electoral votes for each state he carries.[14] Individual voting power within each state is equal, because (we assume) each district has an equal number of voters. All districts have equal voting power in the Electoral College, because they have equal weight, i.e., a single electoral vote; and all states have equal voting power in the Electoral College, because they have equal weight, i.e., two electoral votes. But individual voting power across states is not equal, because districts in different states have different numbers of voters (because House seats must be apportioned in whole numbers) and states with different populations (and numbers of voters) have equal electoral votes. As in the previous discussions, doubly decisive votes can be cast in three distinct contingencies: (1) a vote is decisive in both the voter's district and state and the combined three electoral votes are decisive in the Electoral College; (2) a vote is decisive in the voter's state and the state's two electoral votes are decisive in the Electoral College; and (3) a vote is decisive in the voter's district and the district's one electoral vote is decisive in the Electoral College.

The logic of the Modified District Plan is more complicated than it may at first appear. Because each individual vote counts in two ways, there are logical interdependencies in the way in which district and state electoral votes may be cast. Whichever candidate wins the statewide popular vote *must also* win at least one district electoral vote but, at the same time, *need not* win more than one. Put otherwise, any statewide winner must win at least three of the state's electoral votes but need not win more than that. It follows that the three electoral votes cast by the smallest states are always undivided, just as under the existing 'winner-take-all' Electoral College. In states with four electoral votes, the state popular vote winner is guaranteed a majority of the state's electoral votes (i.e., at least three, with an even split precluded). In states with five electoral votes, the state popular vote winner is guaranteed majority of electoral votes. But in states with six electoral votes, the state popular vote winner may do no better than an even split and, in states with seven or more electoral votes, the state popular vote winner may win fewer than half of

[14] This system is used at present by Maine (since 1972) and Nebraska (since 1992). The 2008 election for the first time produced a split electoral vote in one of these states, namely Nebraska, where Obama carried one Congressional District. A proposed constitutional amendment (the Mundt-Coudert Plan) in the 1950s would have mandated the Modified District Plan for all states.

them—that is, 'election inversions' may occur at the state, as well as the national, level.

However, the preceding remarks pertain only to logical possibilities. Probabilistically, the casting of district and statewide electoral votes is to some degree aligned in random elections (and more so in actual ones). Given that a candidate wins a given district, the probability that the candidate also wins statewide is greater than 0.5—that is to say, even though individual voters cast statistically independent votes, the fact that they are casting individual votes that count in the same way at two levels (district and state) induces a correlation between popular votes at the district and state levels within the same state. As we have seen, this correlation is perfect in the states with only three electoral votes and diminishes as a state's number of electoral votes increases. This implies that the Modified District Plan enhances individual voting power in small states even more than the Pure District Plan does.

Again we follow the procedure outlined earlier. In this case, I generated a sample of 1,080,000 random elections, each with about 122 million voters, in which electoral votes were awarded to the candidates on the basis of the Modified District Plan.[15] For each state, a crosstabulation was constructed and the relevant second-tier probabilities inferred.[16] While the probability of each contingency (straightforwardly calculated) varies considerably with the size of the state, it turns out that the probabilities of double decisiveness in each contingency are essentially constant regardless of state size—namely about 0.0736 in Contingency 1, 0.0502 in Contingency 2, and 0.0253 in Contingency 3—because the same number of electoral votes (three, two, or one, respectively) are at stake regardless of the size of the state.

Figure 10a shows individual voting power across the states under the Modified District Plan.[17] This chart invites comparison with Fig. 10b, showing individual

[15] Again these simulations were generated at the level of the 436 districts, not individual voters. For each random election, the popular vote for one candidate was generated in each Congressional District by drawing a random number from a normal distribution with a mean of $n/2$ and a standard deviation of $\sqrt{.125n}$, where n is the number of voters in the district, i.e., the normal approximation to the Bernoulli distribution with $p = 0.5$. The winner in each district was determined, the district votes in each state were added up to determine the state winner, and electoral votes were allocated accordingly.

[16] Even given this very large sample of elections, the large electorate size meant that few elections were tied at the district or state level, so the relevant electoral vote distributions were taken from a somewhat wider band of elections, in this case those falling within about 0.1 standard deviations of an exact tie.

[17] Unlike those in Fig. 10b, the plotted points in Fig. 10a are subject to some sampling error (though its effects are probably almost invisible), as well as errors due to other approximations noted in the text. However, the most prominent apparent anomalies in Fig. 10a, where voters in a slightly more populous state (e.g., Rhode Island or Iowa) may have somewhat greater voting power than voters is slightly less populous states (e.g., Montana or Kansas) primarily reflect real discrepancies affecting voters in states with approximately similar populations that happen to fall on opposite sides of a threshold in the (whole-number) apportionment of electoral votes. For example, Rhode Island is the smallest state with four electoral votes, while Montana is the largest state with three electoral votes. (Such discrepancies are found in all Electoral College variants that apportion electoral votes into whole numbers.)

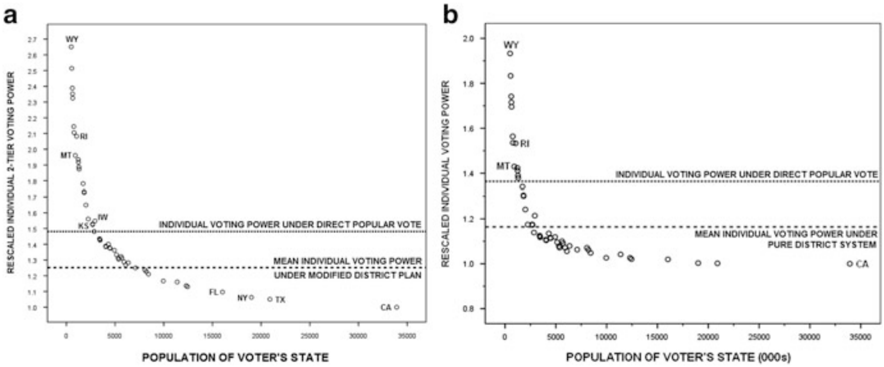


Fig. 10 Individual voting power by state population. (**a**) Under the modified district plan. (**b**) Under the pure proportional plan

voting power by state population under the Pure District Plan. It can be seen that, as anticipated, the winner-take-all effect for three-electorate vote states and the 'winner-take-most' effect for other small-electoral vote states under the Modified District Plan further enhances the voting power of voters in these small states relative to that under the Pure District Plan. In addition, states that are relatively small but not among the smallest (with a population of about 2.5–5 million) are more favored relative to both the smallest states and larger states under the Modified District Plan than the Pure District Plan. Put otherwise, the implicit "voting power by state population curve" in Fig. 10a bends less abruptly in the vicinity of the "southwest" corner of the chart than in Fig. 10b.

Figure 10a also invites comparison with Fig. 2 showing individual voting power under the Modified District plan when calculated in the Banzhaf/Edelman manner. While inequality in voting power is slightly less in Fig. 10a, the main difference is that the (absolute and not rescaled) voting power of all voters is substantially less in Fig. 10a than in Fig. 2, as is indicated by the position of the lines showing (rescaled) individual voting power under direct popular vote. Figure 11 depicts this more directly, by overlaying the two scattergrams and showing absolute, not relative, voting power on the vertical axis. Indeed, we can readily get a good approximation of individual voting power under the Modified District Plan by using the (more straightforward but in principle incorrect) Banzhaf–Edelman mode of calculation in the first instance and then reducing each value by about 20 %. With this correction factor added, Edelman's (2004) conjecture that, with a large number of voters (and states and districts), voting power under the Modified District Plan may be just about the same as when individuals cast two separate and independent votes is sustained.

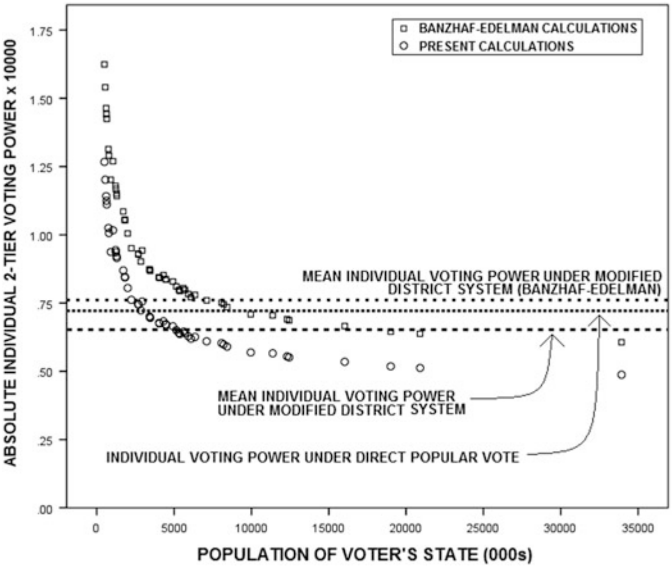


Fig. 11 Present and Banzhaf–Edelman calculations compared for modified district plan

7 Summary and Conclusions

When we try to measure the *a priori* voting power of individual voters under proposed variants of the two-tier U.S. Electoral College system, two plans present special difficulties: the ‘Modified District Plan,’ under which a candidate is awarded one electoral vote for each Congressional District he carries and two electoral votes for each state he carries, and the ‘National Bonus Plan,’ under which a candidate is awarded all the electoral votes of each state he carries (as at present) plus a ‘national bonus’ of some fixed number of electoral votes if he wins the national popular vote. This difficulty arises because, under these arrangements, each voter casts a single vote that counts in two ways: in the voter’s district and state under the Modified District Plan, and in the voter’s state and the nation as a whole under the National Bonus Plan.

In his original analysis of voting power under Electoral College variants, Banzhaf (1968) evaluated voting power under the Modified District Plan by calculating a voter’s two-stage voting power first through the district vote and then through the state vote and then adding the two values together. Unfortunately, this approach cannot be justified, because it ignores interdependencies in the way district and state electoral votes may be cast—in particular, while individuals are casting statistically independent votes, the fact that each is casting a vote that counts in two different upper tiers induces a correlation between popular votes at different levels. That this problem is serious is indicated by the fact that mean individual voting power under the Modified District system, when calculated in the Banzhaf manner, exceeds individual voting power under direct national popular vote, which Felsenthal and Machover (1998) show is a logical impossibility for a simple voting game.

While an analytic solution to this problem may be possible, the difficulties appear to be formidable. Instead, I have proceeded computationally by generating very large samples of random elections, with electoral votes awarded to the candidates on the basis of each plan. This generates a database that can be manipulated to determine the expected distributions of electoral votes for a candidate under specified contingencies with respect to first-tier voting, from which relevant second-tier probabilities can be inferred.

We conclude that the Banzhaf–Edelman calculations get the relative voting power of individual voters just about right but considerably overestimate their absolute voting power.

Acknowledgements This is a revised and expanded version of a paper on "Voting Power with District Plus At-Large Representation" presented at the 2008 Annual Meeting of the (U.S.) Public Choice Society, San Antonio, March 6–9, 2008. It also draws on material from "*A Priori* Voting Power and the U.S. Electoral College" *Homo Oeconomicus*, 26 (3/4, 2009): 341–380. I thank Dan Felsenthal, Moshé Machover, and especially Claus Beisbart for very helpful criticisms and suggestions on early stages of this work.

References

Banzhaf, J. F., III. (1968). One man, 3.312 votes: A mathematical analysis of the Electoral College. *Villanova Law Review, 13*, 304–332.

Beisbart, C. (2007). *One man, several votes*. University of Dortmund.

Beisbart, C., & Bovens, L. (2008). A power measure analysis of the amendment 36 in Colorado. *Public Choice, 134*, 231–246.

Edelman, P. H. (2004). Voting power and at-large representation. *Mathematical Social Sciences, 47*, 219–232.

Feix, M. R., Lepelly, D., Merlin, V. R., & Rouet, J.-L. (2004). The probability of conflicts in a U.S. Presidential type election. *Economic Theory, 23*, 227–257.

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

Miller, N. R. (2009). *A priori* voting power and the U.S. Electoral College. *Homo Oeconomicus, 26*, 341–380.

Miller, N. R. (2011). Voting power in the U.S. Electoral College. In K. Dowding (Ed.), *Encyclopedia of power* (pp. 679–682). Thousand Oaks: Sage Publications.

Miller, N. R. (2012). Election inversions by the U.S. Electoral College. In D. S. Felsenthal & M. Machover (Eds.), *Electoral systems: Paradoxes, assumptions, and procedures* (pp. 93–127). Berlin: Springer.

Schlesinger, A. M., Jr. (2000, December 19). Fixing the Electoral College. *Washington Post*, A39.

Part IV
Penrose, Banzhaf, Shapley–Shubik, et al.

Aspects of Power Overlooked by Power Indices

Manfred J. Holler and Hannu Nurmi

1 Introduction

The background of voting power indices is in game theory and measurement theory. One of their uses is to provide an estimate for the value or payoff that an actor may expect to receive when entering a game. As such they are akin to means, modes and medians. The typical setting of power indices involves evaluation: from a given resource (vote) distribution one aims to estimate the actors' influence over decision outcomes when basically nothing is known about the issues to be decided upon in the game. For example, proportional representation (PR) systems aim at a distribution of parliamentary seats among parties that is nearly identical with the distribution of support given to those parties in the elections. Thereby an (implicit)

This paper was presented at The Leverhulme Trust sponsored Voting Power in Practice Symposium at the London School of Economics, 20–22 March 2011. Comments of the participants, in general, and of Dan S. Felsenthal, in particular, are gratefully acknowledged. We would also like to thank an anonymous referee for several perceptive and constructive suggestions that we have tried to accommodate in the present version.

M.J. Holler (✉)
Department of SocioEconomics, University of Hamburg, Hamburg, Germany

Public Choice Research Centre, University of Turku, Turku, Finland
e-mail: holler@econ.uni-hamburg.de

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

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare, DOI 10.1007/978-3-319-05158-1_12,

assumption is made that distribution of seats coincides with the distribution of legislative influence. This assumption is, however, untenable: x % of seats does not in general give a party x % control over legislation: given majority voting, 51 % of seats imply full control. The a priori voting power indices aim to rectify this by explicitly introducing the decision rule so that it is the decision rule together with the vote distribution that determines the influence over outcomes.

By taking into account the decision rules, the standard a priori voting power indices take a step towards measuring the influence of actors on the decision outcomes. The fact that they sometimes deviate from independent observations about power distribution can partly be explained by their very a priori nature. For example, if an index "assumes" that all coalitions of actors are equally likely, it is to be expected that it provides poor estimates of power distribution in bodies where large classes of coalitions are impossible or extremely rare because of ideological constraints.

The classical power indices have for some time been criticized for ignoring the preferences of actors in coalition formation. In response to this criticism a new type of indices—often called preference-based ones—has been developed (Steunenberg et al. 1999; Napel and Widgrén 2005, 2009). In those indices the power is measured in terms of the distance of outcomes to the actors' ideal points. The main issue in this paper is that the standard a priori voting power indices do what they are supposed to do under very special circumstances only. The same is true—albeit for different reasons—of the preference-based indices.

In the next section the classical indices are introduced and briefly motivated. It is standard to relate them to yes-no decision making. However, in following Sect. 3 it is shown that dichotomous voting typically takes place in a multi-alternative environment, i.e. while the vote is taken between two alternatives at each stage of the procedure, there are several interdependent binary votes in the process. The agenda determines the sequence of these votes. Under certain types of behavioral assumptions the sequence also crucially restricts the feasible outcomes. It is argued that, when compared with marginal changes in voter resource distribution, the control of agenda is of essentially greater importance with regard to the voting outcomes.

Section 4 deals with various monotonicity-related paradoxes in an effort to demonstrate that power under some widely used voting procedures in multiple-alternative settings is not locally monotonic. Hence indices based on this type of monotonicity fail to capture the distribution of power under those procedures. In Sect. 5 we deal with the issue of how voting procedures influence the voting power distribution and whether preference proximity considerations are reconcilable with other intuitively plausible choice principles. In Sect. 6 we turn to paradoxes of composition to illustrate how the very notion of proximity may become ambiguous even in simple game, i.e. dichotomous settings.

2 A Priori Power Indices

The Shapley-Shubik power index (S-S) is a projection of the Shapley value to simple games (Shapley 1953; Shapley and Shubik 1954).[1] It can viewed as a measure based on the assumption that all attitude dimensions (sequences of decision makers in order from the most supportive to the least supportive one) are equiprobable. The two indices named after Penrose and Banzhaf replace this equiprobability of dimensions assumption with one that pertains to actor coalitions (Penrose 1946; Banzhaf 1965). The standardized Penrose-Banzhaf (P-B) index counts for each player the number of winning coalitions where this player has a swing, i.e. where his presence is, *ceteris paribus*, crucial for the coalition to be winning, and divides this number by the sum of swings of all players. The absolute Penrose-Banzhaf index counts the number of swings and divides this by the number of coalitions where the player is present. In contrast to the previous ones, the values of absolute Penrose-Banzhaf index, when summed over the actors, do not in general add up to unity.

In all these three indices the power of a player is determined by the number of winning coalitions in which he is present as an essential member in the sense that should he leave the coalition, it would become non-winning.

Two more recent indices, viz. the public good index (PGI), introduced in Holler (1982), and the Deegan-Packel (1982) index, focus on the minimal winning coalitions, i.e. on coalitions in which all members are decisive in the sense that should *any* one of them leave the coalition, it would become non-winning. The importance of players, and consequently their payoff expectation, is according to the designers of these indices reflected by the number of presences in these types of coalitions.

Table 1 illustrates the above indices in the now bygone EU-15. The differences between the Shapley-Shubik and standardized Penrose-Banzhaf index values are in general very small. The same observation holds for the two indices based on swings in minimal winning coalitions: DP and PGI. Note that countries with larger voting weights have at least as large power values as countries with smaller voting weights.

This monotonicity property, i.e. local monotonicity, is not always satisfied for the values of the Deegan-Packel index and the PGI. The following voting game illustrates this. A voting body consists of 6 persons with voting weights $3, 3, 1, 1, 1, 1$. The decision rule is 6, i.e. any coalition with the sum of voting weights of at least 6 is winning. This yields the PGI value distribution: $5/34, 5/34, 6/34, 6/34, 6/34, 6/34$. In other words, the players with larger voting weights have a smaller PGI value than those with smaller weights. Hence, local monotonicity is violated.

[1]The computation formulae of the indices are listed in Appendix.

Table 1 The Shapley-Shubik (S-S), standardized Penrose-Banzhaf (P-B), Deegan-Packel (DP) and PGI values of countries in the EU-15 for the rule 62/87

Country	No. of votes	S-S index	Std. P-B index	DP index	Holler index
F, G, I, UK	10	0.1167	0.1116	0.0822	0.0809
S	8	0.0955	0.0924	0.0751	0.0743
B, G, N, P	5	0.0552	0.0587	0.0647	0.0650
A, S	4	0.0454	0.0479	0.0608	0.0613
D, Fi, Ir	3	0.0353	0.0359	0.0572	0.0582
L	2	0.0207	0.0226	0.0440	0.0450

3 Agenda-Based Procedures

The simple games are the domain of the above indices of a priori voting power. There are circumstances where simple games are quite natural analysis devices. For example, the votes of confidence or non-confidence in parliamentary systems would seem like simple games in requiring the voters (MPs) to choose one of two exhaustive and mutually exclusive alternatives. Similarly, in most parliaments legislative outcomes are determined on the basis of a binary vote where the winning alternative defeats its competitor in the final contest. Upon closer scrutiny, however, most legislative processes involve more than two decision alternatives. In committee decisions the agenda-building is typically preceded by a discussion in the course of which various parties make proposals for the policy to be taken or candidates for offices. By agenda-based procedures one usually refers to committee procedures where the agenda is explicitly decided upon after the decision alternatives are known. Typical settings of agenda-based procedures are parliaments and committees. One of the crucial determinants of voting power overlooked by power indices is the power of agenda-builder.

Two procedures stand out among the agenda-base systems: (1) the amendment and (2) the successive procedure. Both are widely used in contemporary parliaments. The successive one is based on pairwise comparisons. At each stage of this procedure an alternative is confronted with the set of all remaining alternatives. If it is voted upon by a majority, it is elected and the process is terminated. Otherwise this alternative is set aside and the next one is confronted with all the remaining alternatives. Again the majority decides whether this alternative is elected and the process terminated or whether the next alternative is picked up for the next vote. Eventually one alternative gets the majority support and is elected.

Figure 1 one shows an example of a successive agenda where the order of alternatives to be voted upon is B, A, C, D, E, F and G. Whether this sequence will be followed through depends on the outcomes of the ballots. In general, the maximum number of ballots taken of k alternatives is $k-1$. If an alternative gets a simple majority of votes, it is selected as winner.

The amendment procedure confronts alternatives with each other in pairs so that in each ballot two separate alternatives are compared. Whichever gets the majority

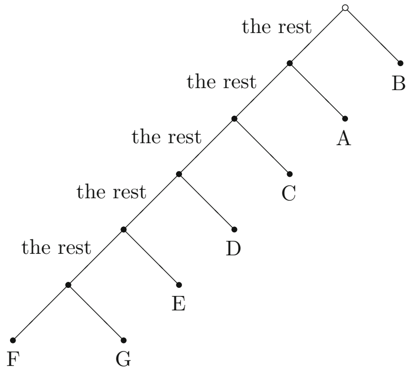


Fig. 1 The successive agenda

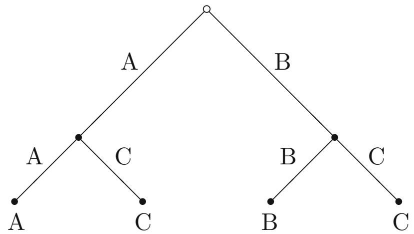


Fig. 2 The amendment agenda

of votes proceeds to the next ballot, while the loser is set aside. Figure 2 shows and example of an amendment agenda over 3 alternatives: A, B and C.

In Fig. 2 alternatives A and B are first compared and the winner is faced with C on the second ballot.

Both the amendment and successive procedure are very agenda-sensitive systems. In other words, two agendas may produce different outcomes even though the underlying preference ranking of voters and their voting behavior remain the same. Under sincere voting—whereby, for all pairs of alternatives A and B, the voter always votes for A if he prefers A to B and vice versa—the well-known Condorcet's paradox provides an example: of the three alternatives any one can be rendered the winner depending on the agenda. To determine the outcomes—even under sincere voting—of successive procedure requires additional assumptions regarding voter preferences over subsets of alternatives. If the voters always vote for the subset of alternatives that contains their first-ranked alternative, the successive procedure is also very vulnerable to agenda-manipulation.

The agenda-based systems have received some attention in the social choice theory. Thus, we know e.g. the following about the amendment and successive systems:

1. Condorcet losers are not elected (not even under sincere voting).
2. Sophisticated voting avoids the worst possible outcomes, i.e those outside the Pareto set.

3. The Condorcet winner is elected (both under sincere and strategic voting) by the amendment procedure.
4. The strong Condorcet winner is elected by both systems.

The first point follows from the observation that the alternative that wins under the amendment procedure has to beat at least one other alternative. Hence, it cannot be the Condorcet loser either. Under the successive procedure if the winner is determined at the final pairwise vote, it cannot be the Condorcet loser. If, on the other hand, the winner appears earlier, it cannot be the Condorcet loser either because it is ranked first by more than half of the voting body.

Sophisticated voting avoids Pareto violations. In other words, if the voters anticipate the outcomes ensuing from various voting strategies, the resulting strategy combinations exclude outcomes for which unanimously preferred outcomes exist (see Miller 1995, p. 87).

That the amendment procedure results in the Condorcet winner under sincere voting, follows from the definition. Finally, the strong Condorcet winner—i.e. one that is ranked first by more than half of the electorate—is elected by both systems regardless of whether the voting is sincere or strategic.

To counterbalance the basically positive results mentioned above, there are some negative ones such as,

1. McKelvey's (1979) results on majority rule and agenda-control.
2. All Condorcet extensions are vulnerable to the no-show paradox (Moulin 1988).
3. Pareto violations are possible under sincere voting.

McKelvey's well-known theorem states that under fairly general conditions—multi-dimensional policy spaces, continuous utilities over the policy space, empty core—*any* alternative can become the voting outcome under amendment procedure if the voters are sincere and myopic. Under these circumstances the agenda-controller determines the outcome even though at every stage of voting the majority determines the winner of the pairwise vote. Although some of the conditions are not so liberal as they seem at first sight, the theorem is certainly important in calling attention to the limits—or rather, lack thereof—that the majority rule *per se* can impose on the possible outcomes. The upshot is that the majority rule guarantees no correspondence between voter opinions and voting outcomes.

Although no analogous result on the outcomes of the successive procedure in multi-dimensional policy spaces exists, it also can be shown to be very vulnerable to agenda-manipulation (Nurmi 2010). In conclusion, then, ignoring the process whereby the sequence of pairwise votes is determined can result in a misleading picture of the influence that various actors exert upon the decision outcomes. Admittedly, the power of the agenda-builder can to some extent be counteracted through sophisticated voting, but even so the best—and in itself exhaustive—characterization of the outcomes reachable by pairwise majority voting, i.e. the Banks set, sometimes leaves a significant maneuvering room for the agenda-builder.

Local monotonicity is a property that many scholars deem particularly important. What it states is that increasing an actor's resources (votes, shares of stock), *ceteris* *paribus*, is never accompanied with a diminution of his voting power. It is known that the Shapley-Shubik and the Penrose-Banhaf indices are locally monotonic, while the indices based on minimal winning coalitions, the Deegan-Packel index and PGI, are not. But is the influence over outcomes always locally monotonic?

4 More Votes, Less Power

The intuitive view of power—voting power included—is based on two tenets:

- the more resources an actor controls, the more often he is on the winning side
- the more powerful an actor, the closer his preferences are to the collective decisions.

Let us look at the former claim first. In voting studies, the resources are typically votes in a voting body. The tenet, thus, has it that the more votes, the more powerful the decision maker. In situations involving more than two alternatives, this tenet has to be essentially qualified, if not downright rejected on the grounds that some widely used voting rules contradict it. In other words, the tenet is at least not universally applicable. In fact, two social choice properties are directly relevant for the rejection of the tenet: non-monotonicity and vulnerability to the no-show paradox. The former means that under some preference profiles it is possible that additional support, *ceteris paribus* would render a winning alternative a non-winning one. On the other hand, systems where some voters might end up with more preferable outcomes by not voting at all than by voting according to their preferences, are vulnerable to the no-show paradox. These two properties are closely related, but not equivalent.

Table 2 illustrates the non-monotonicity of plurality runoff system. Assuming that everyone votes according to his preference, i.e. the voting is sincere, the plurality runoff results in A. Suppose now that the winner had somewhat more support so that two of the voters with $B \succ C \succ A$ ranking had lifted A first, *ceteris paribus*. In this new profile, the runoff would take place between A and C, whereupon C would win. Hence, clearly the $A \succ B \succ C$ group would have done better—been more powerful—with less votes.[2]

Table 3 illustrates a related phenomenon. By abstaining a group of voters may—*ceteris paribus*—improve upon the outcome that would result if they voted according to their preferences. The example is again based on plurality runoff system. With sincere voting, A wins, but if two voters in the $B \succ C \succ A$ group abstain, C wins, an improvement upon A from the view-point of the abstainers.

[2] A referee correctly points out that the outcomes under sincere voting are not Nash equilibria. Indeed, none of the three outcomes is a Nash equilibrium. To wit, if A is the outcome, then Group 2 has an incentive to vote for C at the outset making it thereby the strong Condocet winner and hence the plurality runoff winner as well. The same argument applies *mutatis mutandis* to the two other outcomes B and C.

Table 2 Additional support paradox

22 voters	21 voters	20 voters
A	B	C
B	C	A
C	A	B

Table 3 No-show paradox

5 voters	5 voters	4 voters
A	B	C
B	C	A
C	A	B

Table 4 Schwartz' paradox: an example

Party A 23 seats	Party B 28 seats	Party C 49 seats
a	b	c
b	c	a
c	a	b

Provided that C is closer to the abstainers' preferences than A, the second tenet above is again contradicted.[3]

One more argument can be presented in contradiction to the above tenets. Schwartz (1995) calls it the paradox of representation. But since there are several paradoxes related to representation we shall call it Schwartz' paradox. It is useful to illustrate it in terms of the amendment procedure. Consider Table 4.

Suppose that in parliamentary debate a motion b has been presented and that also an amendment to it c is on the table. Hence we have the amendment agenda:

- motion b vs. amendment c,
- the winner of the preceding vs. a, the status quo

With sincere voting a emerges as the winner. Suppose now that party B would lose all its seats so that parties A and C would share those seats equally. Thus, c would become the (strong) Condorcet winner and hence the winner of the contest here. Again clearly a violation of the tenets above.

The above examples are procedure-related and thus basically avoidable by choosing a monotonic voting system, such as plurality voting or Borda count.

[3]We shall here deal with the general no-show paradox only and omit its strong version. A more comprehensive account of both types is given in Nurmi (2012) which is to a large extent a result of private correspondence with Dan S. Felsenthal dating back to May 2001 and continuing intermittently till early 2011 (Felsenthal 2001–2011).

5 Power and Preference Proximity

Consider a voting body and a very small group of voters with identical preferences in it. Suppose that the voters make a mistake in reporting their preferences in an election. One of the group members may have interpreted the content of decision alternatives incorrectly and the others are following his lead in reporting their preferences in voting. Since we are dealing with a small group of voters, the preference profile containing the intended preferences and the one containing the erroneous preferences should be—if not identical—close to each other. Now, a plausible desideratum for a voting procedure is that mistakes of small voter groups and the accompanying small changes in preference profiles should not result in large changes in ensuing voting outcomes. In particular, the changes in the latter should not be larger as a result of mistaken reports of small voter groups than as a result of mistakes of larger ones. This is intuitively what voting power is about: changing the ballots of big groups should make a larger difference in voting outcomes than changing the ballots of small groups. This *prima facie* plausible desideratum turns, however, out to be incompatible with other intuitively compelling requirements of social choices.

The fundamental results in this area is due to Baigent (1987). To illustrate one of them, consider a drastic simplification of NATO and its policy options with regard to the on-going uprising in Libya.[4] Let us assume that there are only two partners in NATO (1 and 2) and two alternatives: impose a no-fly zone in Libya (NFZ) and refrain from military interference (R) in Libya. To simplify things even further, assume that only strict preferences are possible, i.e both decision makers have a strictly preferred policy. Four profiles are now possible, as shown in Table 5.

We denote the voters' rankings in various profiles by P_{mi} where m denotes the number of the profile and i the voter. We consider two types of metrics: one is defined on pairs of rankings and the other on profiles. The former is denoted by d_r and the latter by d_P. The two metrics are related as follows:

$$d_P(P_m, P_j) = \sum_{i \in N} d_r(P_{mi}, P_{ji}).$$

In other words, the distance between two profiles is the sum of distances between the pairs of rankings of the first, second, etc. voters. No further assumptions on the metric has been made.

Take now two profiles, P_1 and P_3, from Table 5 and express their distance using metric d_P as follows:

$$d_P(P_1, P_3) = d_r(P_{11}, P_{31}) + d_r(P_{12}, P_{32}).$$

[4] The argument is a slight modification of Baigent's (1987, p. 163) illustration.

Table 5 Four two-voter profiles

P_1		P_2		P_3		P_4	
1	2	1	2	1	2	1	2
NFZ	NFZ	R	R	R	NFZ	NFZ	R
R	R	NFZ	NFZ	NFZ	R	R	NFZ

Since, $P_{12} = P_{32} =$ NFZ $\succ$R, and hence the latter summand equals zero, this reduces to:

$$d_P(P_1, P_3) = d_r(P_{11}, P_{31}) = d_r((NFZ \succ R), (R \succ NFZ)).$$

Taking now the distance between P_3 and P_4, we get:

$$d_P(P_3, P_4) = d_r(P_{31}, P_{41}) + d_r(P_{32}, P_{42}).$$

Both summands are equal since by definition:

$$d_r((R \succ NFZ), (NFZ \succ R)) =$$

$$d_r((NFZ \succ R), (R \succ NFZ)).$$

Thus,

$$d_P(P_3, P_4) = 2 \times d_r((NFZ \succ R), (R \succ NFZ)).$$

In terms of d_P, then, P_3 is closer to P_1 than to P_4. This makes sense intuitively.

The proximity of the social choices emerging out of various profiles depends on the choice procedures, denoted by g, being applied. Let us make two very mild restrictions on choice procedures, viz. that they are anonymous and respect unanimity. The former states that the choices are not dependent on the labelling of the voters. The latter, in turn, means that if all voters agree on a preference ranking, then that ranking is chosen. In our example, anonymity requires that whatever is the choice in P_3 is also the choice in P_4 since these two profiles can be reduced to each other by relabelling the voters. Unanimity, in turn, requires that $g(P_1) = NFZ$, while $g(P_2) = R$. Therefore, either $g(P_3) \neq g(P_1)$ or $g(P_3) \neq g(P_2)$. Assume the former. It then follows that $d_r(g(P_3), g(P_1)) > 0$. Recalling the implication of anonymity, we now have:

$$d_r(g(P_3), g(P_1)) > 0 = d_r(g(P_3), g(P_4)).$$

In other words, even though P_3 is closer to P_1 than to P_4, the choice made in P_3 is closer to—indeed identical with—that made in P_4. This argument rests on the assumption that $g(P_3) \neq g(P_1)$. Similar argument can, however, be made for the alternative assumption, viz. that $g(P_3) \neq g(P_2)$. The example, thus, shows that anonymity and respect for unanimity cannot be reconciled with a property called proximity preservation (Baigent 1987; Baigent and Klamler 2004): choices made in profiles more close to each other ought to be closer to each other than those made in profiles less close to each other.

The example shows that small mistakes or errors made by voters are not necessarily accompanied with small changes in voting outcomes. Indeed, if the true preferences of voters are those of P_3, then voter 1's mistaken report of his preferences leads to profile P_1, while both voters' making a mistake leads to P_4. Yet, the outcome ensuing from P_1 is further away from the outcome resulting from P_3 than the outcome that would have resulted had more—indeed both—voters made a mistake (whereupon P_4 would have emerged). This example shows that voter mistakes do make a difference. It should be emphasized that the violation of proximity preservation occurs in a wide variety of voting systems, viz. those that satisfy anonymity and unanimity. This result is not dependent on any particular metric with respect to which the distances between profiles and outcomes are measured. Hence, it applies to all preference-based voting systems.[5] Expressed in another way the result states that in nearly all reasonable voting systems it is possible that a small group of voters has a greater impact on voting outcomes than a big group. Thus, we have yet another way of violating local monotonicity.

6 The Ambiguity of Closeness

Preference-based power measures equate an actor's power with the closeness of the decision outcomes to his ideal point in a policy space.[6] In a single-dimensional policy space, this is a relatively straight-forward matter to determine, but in multidimensional spaces closeness of two points depends on the metric used. With different metrics one may end up with different order of closeness of various points to one's ideal point. But even in cases where the metric is agreed upon, we may encounter difficulty in determining which of two points is closer to an actor's ideal point. Ostrogorski's paradox (Table 6) illustrates this (Rae and Daudt 1976).

There are two decision alternatives, X and Y. An individual decision maker has to choose between them on the basis of information regarding their distance from

[5] A referee correctly points out that we are not requiring that the larger group has identical preferences. Instead their preference changes cancel out each other. Indeed, we have here an instance of reversal bias discussed at some length by Nurmi (2005). The point, however, is that a small group of voters may move the outcome a longer distance than a large—albeit heterogenous—group under specific preference configurations.

[6] This section is based on Nurmi (2010).

Table 6 Ostrogorski's paradox

Issue	Issue 1	Issue 2	Issue 3	Majority alternative
Criterion A	X	X	Y	X
Criterion B	X	Y	X	X
Criterion C	Y	X	X	X
Criterion D	Y	Y	Y	Y
Criterion E	Y	Y	Y	Y

the individual's ideal point on three issues, 1–3. Table 6 indicates which alternative is closer to the voter's ideal point on each criterion and in each issue.[7]

If all issues and criteria are equally important to the individual, it is reasonable to assume that on each criterion the individual prefers that alternative that is closer to his ideal point on more issues than the other alternative. The right-most column indicates these preferred alternatives on each criterion. Under the above assumption of equal importance of criteria and issues, one would expect the individual to choose X rather than Y since X is preferred on three criteria out of five.

However, looking at Table 6 from another angle, it becomes evident that Y should be chosen since on every issue it is the alternative that is closer to the individual's ideal point on a majority of criteria. In other words, there are reasonable grounds for arguing that X is closer to the individual's ideal point than Y, but there are equally strong reasons to make the opposite claim.

Ostrogorski's paradox is one of a larger family of aggregation paradoxes. These play an important role in the social sciences in general and in spatial models in particular. They have, however, less dramatic role in preference-based power indices, since these typically assume away the problem exhibited by the paradox. To wit, it is assumed that the distance measurements are unambiguous—a relatively straight-forward assumption in single-dimensional models—i.e. their approach is to find out power relationships *assuming* that the voters measure distances between alternatives in a given manner. For our purposes Ostrogorski-type paradoxes, however, suggest another overlooked aspect in power studies, viz. the packaging of issues or criteria. This is clearly one facet of the agenda-control problematique that we touched upon earlier. By aggregating or dis-aggregating issues one may change the ordering of alternatives when their closeness determines the choice.

[7] X and Y could be applicants for a job or candidates for a political office. The issues, in turn, could be any three important aspects of the office, e.g. foreign policy, financial policy and education policy. The criteria could be work experience, relevant linguistic skills, relevant formal education, relevant social network and relevant social skills.

7 The Proper Setting for Power Indices

The challenges of a priori voting power indices are mostly related to settings involving more than two alternatives. In decision making involving two alternatives they are still useful tools in assessing the implications of changes in decision rules or seat distributions. The a priori nature should, of course, be held in mind. The practical influence over outcomes may grossly deviate from the a priori index values due to the fact that coalitions tend to have different likelihoods of forming. Also the "nature" of various decision making bodies plays a role in using power indices. Felsenthal and Machover (1998) distinguish between I-power and P-power, while Laruelle and Valenciano (2008) introduce a useful distinction between bargaining and take-it-or-leave-it committees. With these distinctions these authors aim at delineating the conditions of the applicability of the indices. It is likely that further work along this line will follow. Above we argued that agenda-institutions and voting rules deserve attention as determinants of not only voting outcomes but also of the distribution influence among actors. In very general terms, majoritarian voting rules (e.g. amendment, Copeland and Dodgson) assign power to majorities, while positional ones (esp. Borda) assign relatively more power to minorities. When we enter the multiple-alternative environment and leave the simple game setting behind, many kinds of issues arise which always complicate and sometimes contradict the conclusions derived in the two-alternative settings. In the preceding an attempt has been made to examine some of these.

Appendix

The Shapley-Shubik index value of player i is:

$$\phi_i = \Sigma_{S \subseteq N} \frac{(s-1)!(n-s)!}{n!}[v(S) - v(S \setminus \{i\})].$$

Here s denotes the number of members of coalition S and $n!$ is defined as the product $n \cdot (n-1) \cdot (n-2) \cdot \ldots 2 \cdot 1$. The expression in square brackets differs from zero just in case S is winning but $S \setminus \{i\}$ is not. In this case, then, i is a decisive member in S. In other words, i has a swing in S. Indeed, the Shapley-Shubik index value of i indicates the expected share of i's swings in all swings assuming that coalitions are formed sequentially.

Player i's PGI value H_i is computed as follows:

$$H_i = \frac{\Sigma_{S* \subseteq N}[v(S*) - v(S * \setminus \{i\})]}{\Sigma_{j \in N} \Sigma_{S* \subseteq N}[v(S*) - v(S * \setminus \{j\})]}.$$

Here $S*$ is a minimal winning coalition, i.e. every proper subset of $S*$ is a losing coalition.

The Deegan-Packel index value of player i, denoted DP_i, in turn, is obtained as follows:

$$DP_i = \frac{\Sigma_{S*\subseteq N} 1/s[v(S*) - v(S * \backslash\{i\})]}{\Sigma_{j\in N} \Sigma_{S*\subseteq N} 1/s[v(S*) - v(S * \backslash\{j\})]}.$$

The standardized Banzhaf index value of i is defined as:

$$\bar{\beta}_i = \frac{\Sigma_{S\subseteq N}[v(S) - v(S \setminus \{i\})]}{\Sigma_{j\in N} \Sigma_{S\subseteq N}[v(S) - v(S \setminus \{j\})]}.$$

The absolute Penrose-Banzhaf index (Penrose 1946; Banzhaf 1965), in turn, is defined as:

$$\beta_i = \frac{\Sigma_{S\subseteq N}[v(S) - v(S \setminus \{i\})]}{2^{n-1}}.$$

References

Baigent, N. (1987). Preference proximity and anonymous social choice. *The Quarterly Journal of Economics*, *102*, 161–169.

Baigent, N., & Klamler, C. (2004). Transitive closure, proximity and intransitivities. *Economic Theory*, *23*, 175–181.

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

Deegan, J., & Packel, E. W. (1982). To the (minimal winning) victors go the (equally divided) spoils: a new power index for simple n-person games. In S. J. Brams, W. F. Lucas, & P. D. Straffin (Eds.), *Political and related models in applied mathematics*. New York: Springer.

Felsenthal, D. S. (2001–2011). private communications 2001–2011.

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

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

Laruelle, A., & Valenciano, F. (2008). *Voting and collective decision-making. Bargaining and power*. Cambridge: Cambridge University Press.

McKelvey, R. (1979). General conditions for global intransitivities in formal voting models. *Econometrica*. *47*, 1085–1112.

Miller, N. (1995). *Committees, agendas, and voting*. Chur: Harwood Academic Publishers.

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

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

Napel, S., & Widgrén, M. (2009). Strategic a priori power in the European Union's codecision procedure. *Homo Oeconomicus*, *26*, 297–316.

Nurmi, H. (2005). A responsive voting system. *Economics of Governance*, *6*, 63–74.

Nurmi, H. (2010). On the relative unimportance of voting weights: observations on agenda-based voting procedures. In M. Cichocki, & K. Życzkowski (Eds.), *Institutional design and voting power in the European Union* (pp. 171–180). Farnham, Surrey: Ashgate.

Nurmi, H. (2012). On the relevance of theoretical results to voting system choice. In D. S. Felsenthal & M. Machover (Eds.), *Electoral systems: Paradoxes, assumptions, and procedures* (pp. 255–274). Berlin: Springer.

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

Rae, D., & Daudt, H. (1976). The ostrogorski paradox: A peculiarity of compound majoroty decision. *European Journal of Political Research 4*, 391–398.

Shapley, L. S. (1953). A value for n-person games. *Annals of Mathematical Studies*, *28*, 307–317.

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

Schwartz, T. (1995). The paradox of representation. *The Journal of Politics*, *57*, 309–323.

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

Banzhaf–Coleman and Shapley–Shubik Indices in Games with a Coalition Structure: A Special Case Study

Maria Ekes

1 Introduction

In the paper we investigate how to measure the power of individuals in a voting body possibly divided into some parties. We are modeling such situation in two different ways: by applying the framework of games with a priori unions (Owen 1977) and by applying composite games (Felsenthal and Machover 1998). In both cases we measure the power of individual voters using Shapley–Shubik and Banzhaf–Coleman indices. We make simulations for a specific voting body composed of 100 members and we compare both approaches. The aim of the paper is to compare the behavior of both indices in those frameworks and to find similarities and differences between them, implied by changes of the size and composition of coalition structures as well as by different methodology of measuring the voters' power (composite game versus game with a priori unions).

We begin with describing the formal model. In the sequel we present the results of simulations for a voting body composed of 100 members with various divisions into parties.

This paper was presented to The Leverhulme Trust sponsored Voting Power in Practice Symposium at the London School of Economics, 20–22 March 2011.

M. Ekes (✉)
Warsaw School of Economics, Warsaw, Poland
e-mail: maria.ekes@gmail.com

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare,
DOI 10.1007/978-3-319-05158-1_13,

2 Model

Let $N = \{1,\ 2\ \ldots\ n\}$ denote the set of voters (or seats). We consider a decision-making situation in which the voting body is supposed to make a decision (to pass or to reject a proposal) by means of a voting rule. We assume that voters who do not vote for a proposal (do not vote "yes") vote against it and there is no possibility of abstention. The voting rule specifies whether the set of voters who accepted the proposal forms a winning coalition or not. Formally, we have 2^n possible coalitions (vote configurations) $S \subseteq N$. The voting rule is then defined by the set of winning coalitions W. Usually it is assumed that

- $\varnothing \notin W$,
- $N \in W$,
- If $S \in W$ then $N - S \notin W$,
- If $S \in W$ and $S \subset T$ then $T \in W$.

The voting rule is equivalently given by a simple voting game v_W as follows

$$v_W(S) = \begin{cases} 1 \text{ if } & S \in W \\ 0 \text{ if } & S \notin W \end{cases}$$

for each $S \subseteq N$.

We say that a voter i is *critical* for a coalition S if $v_W(S) = 0$ and $v_W(S \cup \{i\}) = 1$ or $v_W(S) = 1$ and $v_W(S \backslash \{i\}) = 0$.

The Banzhaf–Coleman index of a voter j in this framework is the probability of a voter to be critical assuming that all voting configurations are equally probable, that is

$$\beta_j(W) = \frac{\#\{S \subset N : (j \in S \in W \wedge S - \{j\} \notin W) \vee (j \notin S \notin W \wedge S \cup \{j\} \in W)\}}{2^n}$$
$$= \frac{1}{2^{n-1}} \sum_{\substack{S \subset N \\ j \in S}} \Big(v_W(S) - v_W\Big(S - \{j\}\Big)\Big).$$

The Shapley–Shubik index of a voter j is a truncation of the Shapley value defined for simple games and it is given by the formula

$$Sh_j(W) = \sum_{\substack{S \subset N \\ j \in S}} \frac{(s-1)!\,(n-s)!}{n!} \left(v_W(S) - v_W\left(S - \{j\}\right)\right),$$

where $s = |S|$. Shapley–Shubik index has also a probabilistic interpretation–if we assume that all orderings of voters are equally probable, then the Shapley–Shubik index of a voter j is the probability, that this voter is pivotal (i.e. changes the already existing coalition from loosing to winning, regardless what happens after his accession). Most important characterizations of BC and SS indices are given in Banzhaf (1965), Coleman (1964), Dubey (1975), Owen (1978), Penrose (1946), Shapley (1953), and Shapley and Shubik (1954).

In real world voting bodies the situation is more complicated since voters are divided into some parties ex ante, which may constrain the actual voting behavior. This partition may be the consequence of the political party membership, which is an obvious reason of some constraints in voting in bodies like parliaments. It might also reflect different national interests of citizens of various members of international communities like EU or IMF. This situation can be described by games with a priori unions (precoalitions) introduced by Owen (1977). Let $T = (T_1, T_2, \ldots, T_m)$ be a partition of the set N into subsets which are nonempty, pairwise disjoint and $\bigcup_{i=1}^{m} T_i = N$. The sets T_i are called precoalitions (a priori unions) and they can be interpreted as parties occupying seats in the voting body (note that some of T_i can be singletons). Let M denote the set of all precoalitions, that is $M = \{1, 2, \ldots, m\}$. Owen proposed the modification of both Shapley value (Owen 1977) and Banzhaf–Coleman index (Owen 1981) for games with a priori unions and we will be dealing in this paper with these modifications. The formulae are as follows:

– Modification of Banzhaf–Coleman index–for a voter j in a union T_i we have

$$O_j^{BC}(W,T) = \frac{1}{2^{m+t_i-2}} \sum_{Q \subset M-\{i\}} \sum_{\substack{K \subset T_i \\ j \in K}} v_w(N(Q) \cup K) - v_w(N(Q) \cup (K - \{j\})),$$

where $N(Q) = \bigcup_{p \in Q} T_p$ and t_i denotes the cardinality of T_i;

– Modification of the Shapley value (or Shapley–Shubik index)–for a voter j in a union T_i we have

$$O_j^{SS}(W,T) = \sum_{\substack{H \subset M \\ i \notin H}} \sum_{\substack{S \subset T_i \\ j \notin S}} \frac{h!\,(m-h-1)!s!\,(t_i-s-1)!}{m!t_i!} (v_W(H \cup S \cup \{j\})) - (v_W(H \cup S)),$$

where h denotes the cardinality of the set H, t_i denotes the cardinality of the party T_i and s denotes the cardinality of the coalition S. This index is often called Owen index. Since we refer here to both modified indices–Banzhaf–Coleman's and Shapley–Shubik's–we shall use the term "coalitional index" in order to avoid misunderstanding.

The coalitional BC index is the ratio of the number of coalitions for which the voter $j \in T_i$ is critical and no coalition different from T_i can be broken to the total number of such coalitions. Laruelle and Valenciano (2004) have given three different probabilistic interpretation of the modified BC index. The coalitional BC index was axiomatized by Albizuri (2000).

When calculating coalitional SS index of a player $j \in T_i$ we restrict the number of possible permutations of the set of players. We take into account only those permutations in which all players from each precoalition appear together. To find all such permutations we need to order the parties first and then to order the players in each party. Coalitional SS index can therefore be interpreted as a probability of a player $j \in T_i$ being pivotal provided that all permutations of the set of players, respecting the coalition structure, are equally probable. Coalitional SS index was also axiomatized in various ways; see e.g. Owen (1977) or Hart and Kurz (1983).

There is also another possibility of measuring the decisiveness of each voter in the context of games with an a priori coalition structure. Suppose that within each party the proposal is accepted or rejected by simple majority voting and then all members of the party vote according to the decision made by inside party voting. This is the case of a composite game (see Felsenthal and Machover 1998). In this case the BC index of a member of a party T_i is the product of his index in the simple majority voting game inside the party and the index of the party T_i treated as a player in the top game. In that game the set of players is M, that is players are parties and the set of winning coalitions is $W_T = \{Q \subset M : N(Q) \in W\}$, so for a voter $j \in T_i$ we have

$$\beta_j^c (W, T) = \beta_j (W_{t_i}) \cdot \beta_i (W_T),$$

where $W_{t_i} = \{K \subset T_i : \#K \geq [\frac{t_i}{2}] + 1\}$ and the symbol $[x]$ denotes the largest integer not greater than x, for any real x. In fact this is the BC index in the composite game with the top W_T and the components W_{T_i} for $i = 1, 2, \ldots, m$.

SS index in a composite game (we will denote it by $Sh^c(W, T)$) does not have such "product" property–we calculate it directly from its definition.

In the sequel we present an example of a voting body composed of 100 voters, who are divided into two parties or vote independently. We calculate the power of voters for both approaches (game with precoalitions and composite game) and using both indices–BC and SS–for all possible configurations of sizes of parties. A part of the results presented here is also examined in Ekes (2006).

3 Description of the Special Case

We consider the situation where the voting body is composed of 100 voters, who are members of one of two existing parties or who are voting as independent voters. The coalition structure is therefore the following: $T = (T_1, T_2, \{j_1\}, \ldots, \{j_l\})$, where $2 \leq t_1, t_2$ and $t_1 + t_2 + l = 100$. We assume that the voting rule in our example is the simple majority, which means that any proposal is accepted if it has at least 51 votes for. We are not interested in case where a single party constitutes the winning majority, therefore we assume that $t_1, t_2 \leq 50$. We have calculated values of all four indices: $O^{BC}(T)$, $O^{SS}(T)$, $\beta^c(T)$ and $Sh^c(T)$ for all possible configurations of sizes of parties and for all voters (we omit the symbol W in the notation of indices since the simple majority rule defines the set of winning coalitions in the game with precoalitions as well as in the composite game). We also note that due to the symmetry of SS and BC indices, the power of all voters belonging to the same party is equal and the power of all independent voters is the same. Therefore we will use the notation $O^{BC}_{T_i}(T)$, $O^{SS}_{T_i}(T)$, $\beta^c_{T_i}(T)$, $Sh^c_{T_i}(T)$ for $i = 1, 2$ and $O^{BC}_{j_k}(T)$, $O^{SS}_{j_k}(T)$, $\beta^c_{j_k}(T)$, $Sh^c_{j_k}(T)$ for $k = 1,2,\ldots, l$. Parties are symmetric in our case.

If we consider the coalition structures of the form $T^1 = \{T^1_1, T^1_2, \{j^1_1\}, \ldots, \{j^1_l\}\}$, $T^2 = \{T^2_1, T^2_2, \{j^2_1\}, \ldots, \{j^2_l\}\}$ such that $t^2_1 = t^1_2 \wedge t^1_1 = t^2_2$, then the value of all considered indices of the first party members given the coalition structure T^1 is equal to the value of respective indices of the second party members given the coalition structure T^2 while the power of independent voters measured by any of considered indices is the same for both coalition structures T^1 and T^2. This observation allows considering only the value of all indices for members of the first party and for independent voters. The number of elements of our coalition structure is equal to $100 - t_1 - t_2 + 2 = l + 2$.

Let us introduce an additional notation. In order to calculate SS index of a T_1 member in the composite game we have two find the set of all coalitions for which the first party is decisive in the weighted voting game of parties (by parties we mean the two "large" parties and all independent voters). We denote this set of coalitions by $Dec(T_1)$. For a coalition $C \in Dec(T_1)$ we find the number of independent players in this coalition and we denote it by $l(C)$. And finally we take $\tau_1 = [\frac{t_1}{2}]$, $\tau_2 = [\frac{t_2}{2}]$. In all formulae below we will assume that $\binom{n}{k} = 0$ for $n <$ k. Therefore in a composite game we have:

- The BC index of a member of T_1 is given by:

$$\beta^c_{T_1}(T) = \frac{1}{2^{100-t_2}} \binom{t_1 - 1}{\tau_1} \left(\sum_{s=\max(51-t_1-t_2,0)}^{\min(50-t_2,l)} \binom{l}{s} + \sum_{s=51-t_1}^{\min(l,50)} \binom{l}{s} \right),$$

  (we assume that $t_1, t_2 \le 50$);
- The SS index of a member of T_1 is given by:

$$Sh^c_{T_1}(T) = \overline{P}_1 + \overline{P}_2,$$

  where

$$\overline{P}_1 = \frac{1}{100!} \sum_{\substack{C \in Dec(T_1) \\ T_2 \notin C}} \sum_{p_2=0}^{\tau_2} \binom{t_2}{p_2} \binom{t_1 - 1}{\tau_1} \binom{l}{l(C)}$$
$$\times (p_2 + l(C) + \tau_1)!\,(100 - (p_2 + l(C) + \tau_1) - 1)!$$

$$\overline{P}_2 = \frac{1}{100!} \sum_{\substack{C \in Dec\,(T_1) \\ T_2 \in C}} \sum_{p_2=\tau_2+1}^{t_2} \binom{t_2}{p_2} \binom{t_1-1}{\tau_1} \binom{l}{l(C)}$$

$$\times (p_2 + l(C) + \tau_1)!\,(100 - (p_2 + l(C) + \tau_1) - 1)!$$

- The BC index of an independent voter j is equal to:

$$\beta^c_{j_k}(T) = \frac{1}{2^{l+1}} \left(\binom{l-1}{50} + \binom{l-1}{50-t_1} + \binom{l-1}{50-t_2} + b \right),$$

where $b = \binom{l-1}{50-t_1-t_2}$ if $t_1 + t_2 \leq 50$ and $b = 0$ otherwise;
- The SS index of an independent voter j_k is equal to:

$$Sh^c_{j_k}(T) = \overline{S}_1 + \overline{S}_2 + \overline{S}_3 + \overline{S}_4,$$

where

$$\overline{S}_1 = \frac{1}{100!} \binom{l-1}{50} \sum_{p_1=0}^{\tau_1} \sum_{p_2=0}^{\tau_2} \binom{t_1}{p_1} \binom{t_2}{p_2}$$

$$\times (50 + p_1 + p_2)!\,(100 - (50 + p_1 + p_2) - 1)!,$$

if $l > 50$, otherwise $\overline{S}_1 = 0$,

$$\overline{S}_2 = \frac{1}{100!} \binom{l-1}{50-t_1} \sum_{p_1=\tau_1+1}^{t_1} \sum_{p_2=0}^{\tau_2} \binom{t_1}{p_1} \binom{t_2}{p_2}$$

$$\times (50 - t_1 + p_1 + p_2)!\,(100 - (50 - t_1 + p_1 + p_2) - 1)!,$$

if $t_2 \neq 50$, otherwise $\overline{S}_2 = 0$,

$$\overline{S}_3 = \frac{1}{100!} \binom{l-1}{50-t_2} \sum_{p_1=0}^{\tau_1} \sum_{p_2=\tau_2+1}^{t_2} \binom{t_1}{p_1} \binom{t_2}{p_2}$$

$$\times (50 - t_2 + p_1 + p_2)!\,(100 - (50 - t_2 + p_1 + p_2) - 1)!,$$

if $t_1 \neq 50$, otherwise $\overline{S}_3 = 0$,

$$\overline{S}_4 = \frac{1}{100!}\binom{l-1}{50-t_1-t_2}\sum_{p_1=\tau_1+1}^{t_1}\sum_{p_2=\tau_2+1}^{t_2}\binom{t_1}{p_1}\binom{t_2}{p_2}$$
$$\times (50-t_1-t_2+p_1+p_2)!\,(100-(50-t_1-t_2+p_1+p_2)-1)!,$$

if $t_1 + t_2 \le 50$, otherwise $\overline{S}_4 = 0$.

If we consider a game with precoalitions, then Owen modifications of concerned indices are calculated using the following formulae:

– The O^{BC} index of a member of T_1 is equal to:

$$\frac{1}{2^{t_1+l}}\binom{t_1-1}{50-t_2} + \sum_{s=1}^{\min(50,l)}\binom{l}{s}\binom{t_1-1}{50-s} + \sum_{s=1}^{\min(50-t_2,l)}\binom{l}{s}\binom{t_1-1}{50-t_2-s};$$

– The O^{SS} index of a member of T_1 is equal to:

$$O_{T_1}(T) = P_1 + P_2,$$

where

$$P_1 = \sum_{s=51-t_1}^{\min(50,l)}\binom{l}{s}\binom{t_1-1}{50-s}\frac{s!\,(l+1-s)!\,(50-s)!\,(t_1-(50-s)-1)!}{t_1!\,(l+2)!},$$

if $t_1 + l \ge 51$, otherwise $P_1 = 0$ and

$$P_2 = \sum_{s=\max(51-t_1-t_2,0)}^{50-t_2}\binom{l}{s}\binom{t_1-1}{50-t_2-s}$$
$$\times \frac{(1+s)!\,(l-s)!\,(50-t_2-s)!\,(t_1-(50-t_2-s)-1)!}{t_1!\,(l+2)!}.$$

– The O^{SS} index of an independent voter j_k is equal to:

$$O_{j_k}(T) = S_1 + S_2 + S_3 + S_4,$$

where

$$S_1 = \binom{l-1}{50}\frac{50!\,(l+1-50)!}{(l+2)!}, \text{ if } l > 50, \text{ otherwise } S_1 = 0$$
$$S_2 = \binom{l-1}{50-t_1}\frac{(50-t_1+1)!\,(l+1-(50-t_1+1))!}{(l+2)!},$$

if $t_2 \ne 50$, otherwise $S_2 = 0$,

$$S_3 = \binom{l-1}{50-t_2} \frac{(50-t_2+1)!\,(l+1-(50-t_2+1))!}{(l+2)!},$$

if $t_1 \neq 50$, otherwise $S_3 = 0$,

$$S_4 = \binom{l-1}{50-t_1-t_2} \frac{(50-t_1-t_2+2)!\,(l+1-(50-t_1-t_1+2))!}{(l+2)!},$$

if $t_1 + t_2 \leq 50$, otherwise $S_4 = 0$.

We do not present the formula for coalitional BC index of an independent voter in a game with coalition structure because it is equal to his BC index in a composite game. It follows from the fact, that the internal power of a member of a "singleton party" is equal to 1 so $\beta^c_{j_k}(W,T) = \beta_{j_k}(W_T)$ for an independent voter j_k and it is equal to $O^{BC}_{j_k}(W,T)$ since swings of the player j_k (or a party composed only of the player j_k) are exactly the same in both cases.

Another important remark is that coalitional SS index has a product property which is similar to the property of BC index in a composite game. After simplification of the formula for $O^{SS}_{T_1}(T)$ we obtain:

$$O^{SS}_{T_1}(T) = \frac{1}{t_1}\left(\tilde{P}_1 + \tilde{P}_2\right),$$

where

$$\tilde{P}_1 = \sum_{s=51-t_1}^{\min(50,l)} \binom{l}{s} \frac{s!\,(l-s+1)!}{(l+2)!}$$

if $t_1 + l \geq 51$, otherwise $\tilde{P}_1 = 0$ and

$$\tilde{P}_2 = \sum_{s=\max(51-t_1-t_2,0)}^{50-t_2} \binom{l}{s} \frac{(1+s)!\,(l-s)!}{(l+2)!}.$$

This new formula has an interesting interpretation–it is the product of the SS index of a member of a party T_1 in a (arbitrary) majority voting game inside this party and the SS index of this party in a top game among parties, which is the weighted majority voting game with the quota 51.

4 Composite Game: Presentation of Results

We begin the analysis of our simulations with the case of composite game. First we consider the power of the first party members. Figures 1 and 2 present the power of members of the first party as the function of the size of the second party.

What we can observe at those charts is that values of both indices–BC and SS in a composite game–decrease monotonically with the increasing size of the second party. It means that the power (measured by BC or SS) of the first party's member goes down as the size of the opponent increases. We cannot of course compare

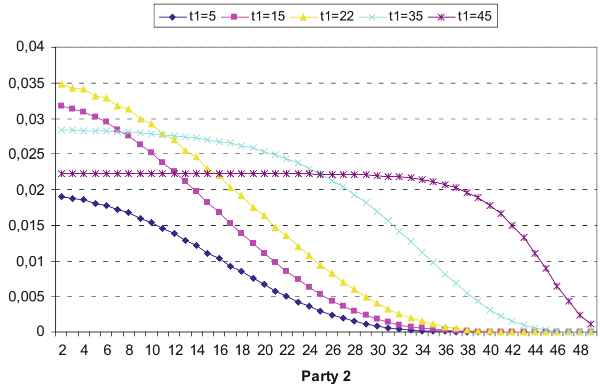


Fig. 1 The SS index of the member of T_1 (as a function of t_2) for various configurations of sizes of both parties

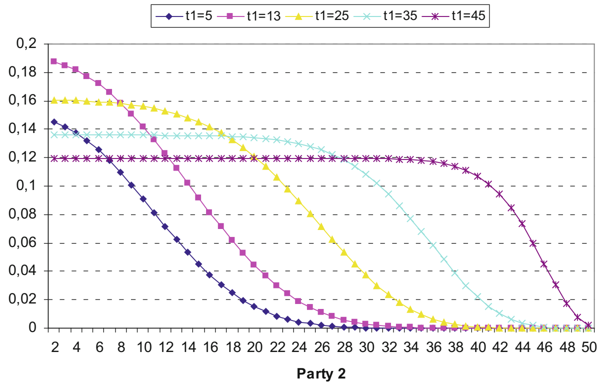


Fig. 2 The BC of the member of T_1 (as a function of t_2) for various configurations of sizes of both parties

values of those indices since one of them is normalized and the other is not, but the shape of curves illustrating changes of values of both indices is very similar. The global maximum of the value of SS index in a composite game (for the member of the party T_1) is attained in the configuration $t_1 = 22$, $t_2 = 2$, while the global maximum of the value of the BC index in a composite game is achieved for $t_1 = 13$, $t_2 = 2$. If the size of the first party increases, the initial value of both indices grows up until t_1 becomes equal to 22 or 13 respectively and then the initial point is coming down. For large values of t_1 the power of the member of the first party measured by both indices is almost constant as a function of t_2, it decreases only for large, almost maximal, sizes of the second party.

We found it interesting to check in what configurations the power of a member of the first party is maximal while the size of the second party is fixed. Both indices have very similar properties also in this case. The point at which the maximal value of the SS index and BC index in a composite game is attained depends on the fixed size of the second party. For $t_2 = 2$ the maximal power of the member of the first party measured by the SS index is attained for $t_1 = 22$ and maximal power measured by BC index is achieved for $t_1 = 13$. If we increase the fixed t_2 , then the value of t_1 at which the maximum of each index is achieved also increases. For $t_2 > 43$ maximum

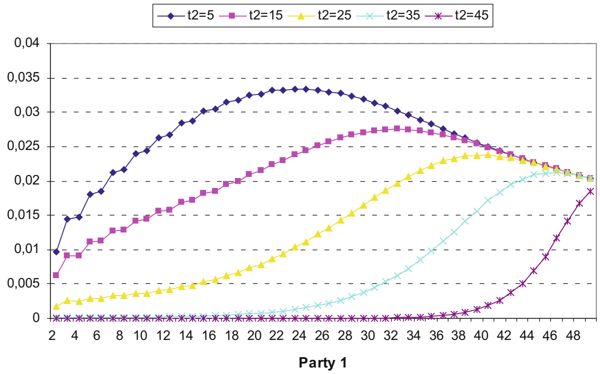


Fig. 3 The SS index of the member of T_1 (as a function of t_1) for various configurations of sizes of both parties

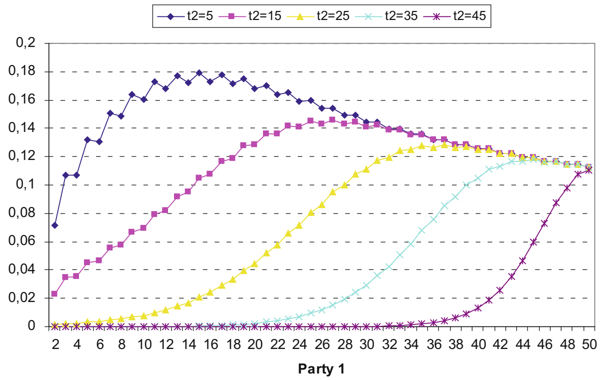


Fig. 4 The BC index of the member of T_1 (as a function of t_1) for various configurations of sizes of both parties

of power of the party T_1 members is achieved in the situation where the first party is of maximal size for both SS and BC index.

In the sequel we comment charts illustrating the behavior of SS and BC indices treated as functions of the own party's size for fixed values of t_2 (Figs. 3 and 4).

The most striking observation is that the power of the member of T_1 measured by both indices–SS and BC–in a composite game is not an increasing function of the own party's size for most values of t_2. For a fixed size of the opponent, the power of the member of the first party increases, attains the maximum and then decreases with an increasing size of the own party. Only for large sizes of the opponent party, the power is an increasing function of the own size. Again the shape of curves is very similar for both indices. The phenomenon of non-monotonicity of BC index of a party member treated as a function of the own party's size in composite games was also examined in the paper of Leech and Leech (2006), where it was interpreted

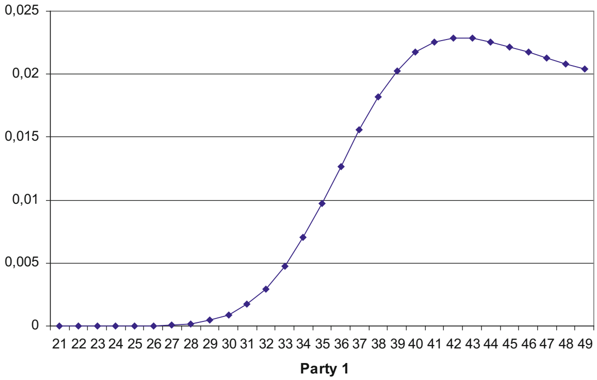


Fig. 5 The SS index of the member of T_1 (as a function of t_1)—the migration from T_2 to T_1 ($l = 30$)

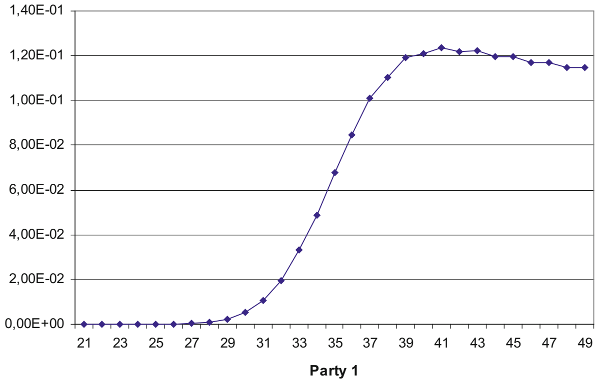


Fig. 6 The BC index of the member of T_1 (as a function of t_1)—the migration from T_2 to T_1 ($l = 30$)

as a tradeoff between the increasing power of a party T_1 as a player in the top game and decreasing power of a party member.

Up to this point we have only considered the migration from one party to the set of independent voters: we have fixed the size of one party and increased the size of another party. Now we take a look to the behavior of both indices if the migration appears between parties.

We assume that members of the second party are joining the first party. Next figures illustrate the influence of this kind of changes in the configuration of sizes on both indices.

Figures 5, 6, 7, and 8 show that SS index and BC index of a first party member is (in most cases) not monotonic with respect to the own party's size in the situation where members of the opponent party are joining T_1. The power of a voter in T_1 measured by BC index grows up, attains its maximum and then decreases, while if we use SS index the situation is different only for large values of l, then the power of T_1 members is an increasing function of t_1.

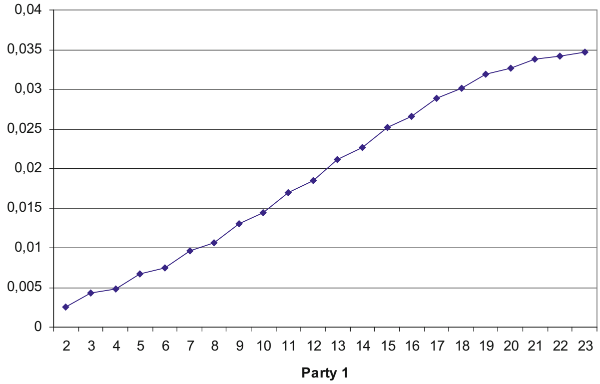


Fig. 7 The SS index of the member of T_1 (as a function of t_1)—the migration from T_2 to T_1 ($l = 75$)

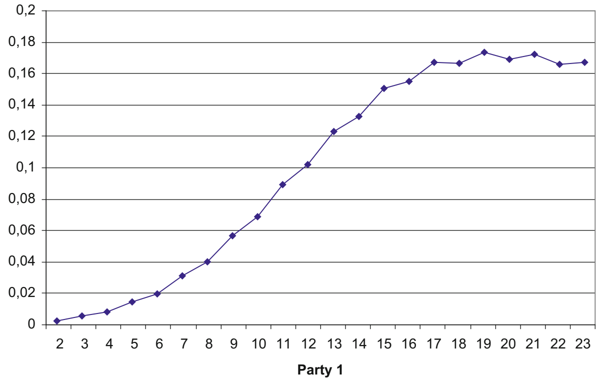


Fig. 8 The BC index of the member of T_1 (as a function of t_1)—the migration from T_2 to T_1 ($l = 75$)

If we compare the power of individual voters in a composite game we also obtain similar results for the SS and BC indices. Below we present some figures showing the value of each of two considered indices in a composite game for various configurations of sizes of both parties.

We treat the power of an independent voter as a function of the size of the first party with the number of all independent voters fixed (if we fix l, then choosing t_1 we determine also t_2). Figures 9, 10, 11, and 12 reflect the fact that the SS index and BC index of an independent voter in a composite game is symmetric with respect to t_1 and t_2. This fact is obvious, because if the size of one party grows up then the size of the second one goes down (the sum $t_1 + t_2$ is fixed). What is more interesting is that maximal power of an independent voter measured in both ways is achieved in the situation where both parties are of the same size, or the difference between their sizes is equal to 1–in this case we have two points with the same maximal value of

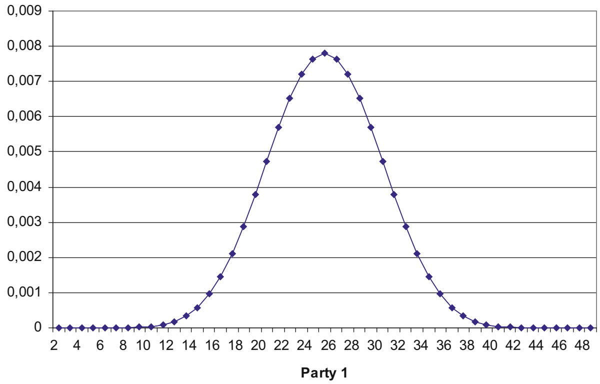


Fig. 9 The SS index of an independent voter as a function of the size of T_1 for $l = 50$

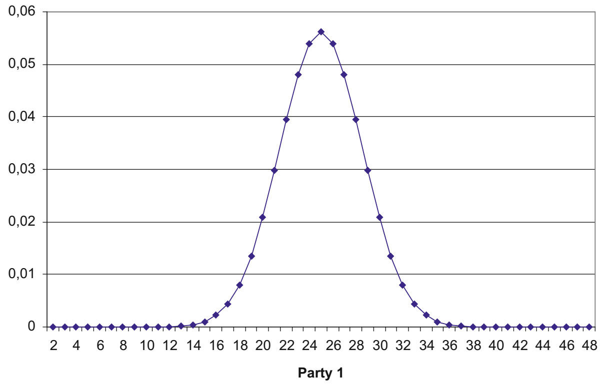


Fig. 10 The BC index of an independent voter as a function of the size of T_1 for $l = 50$

power. The global maximum of the SS index and BC index of an independent voter is attained in case where there is only one such voter.

5 Game with Precoalitions: Presentation of Results

In games with a coalition structure the behavior of coalitional SS index and coalitional BC index is different from the case of composite games. Moreover, both indices differ in their behavior much more than in composite games. We will present the results in the same order as in the previous section.

First we consider the behavior of both indices for members of the party T_1, assuming that the size of the own party is fixed (therefore we treat the power of first party members as the function of the size of the second party).

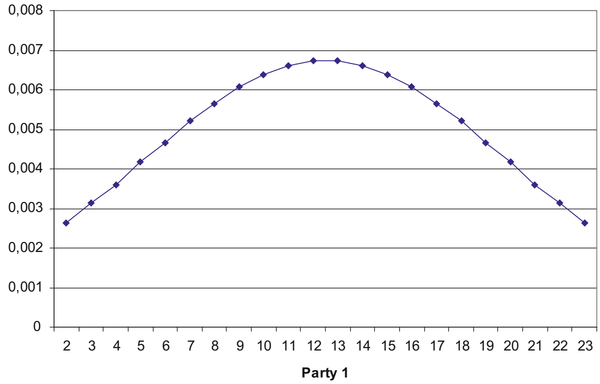


Fig. 11 The SS index of an independent voter as a function of the size of T_1 for $l = 75$

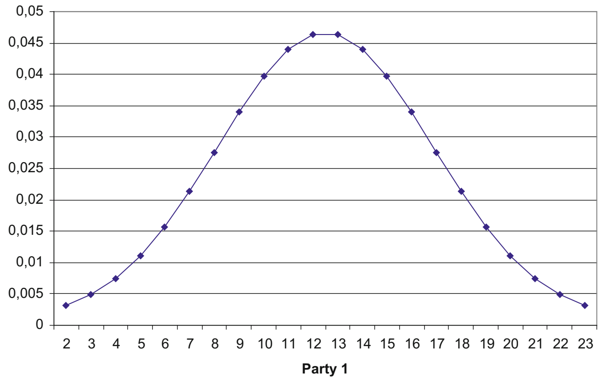


Fig. 12 The BC index of an independent voter as a function of the size of T_1 for $l = 75$

Coalitional SS index of a voter from the party T_1 is a decreasing function of t_2 (for all values of t_1) and attains its maximum in the situation where this party has the maximal possible number of members equal to 50 and the opponent has minimal possible number of members equal to 2 (Fig. 13). Moreover, if we fix the number t_2 , then the maximal value of the coalitional SS index of a voter from the party T_1 is achieved in the situation where the size of the party T_1 is maximal (equal to 50), which means that it does not depend on (fixed) t_2, which was the case in composite games.

The situation appears to be quite different if we consider the coalitional BC index. First note that if we compare the situation where some of the sets T_i are singletons with the situation where singletons join together and form a new party, then the value of the coalitional BC index for members of a new party is the same as it was in the previous partition. Formally, suppose that the partition T is of the form $T = (\{j_1\}, \ldots, \{j_k\}, T_{k+1}, \ldots, T_m)$, where $\# T_i \geq 2$ for $i = k + 1, \ldots, m$ and the

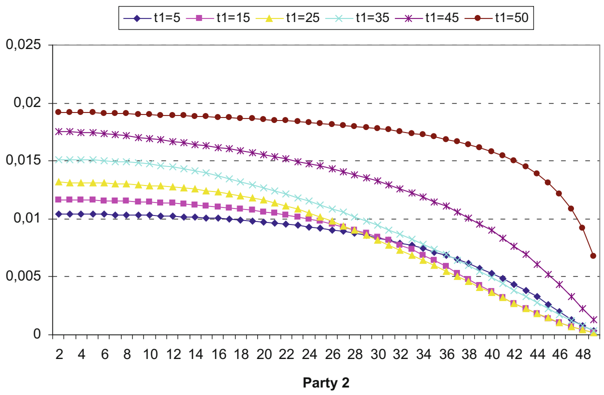


Fig. 13 Coalitional SS index of a member of T_1 (as a function of t_2) for various configurations of sizes of both parties

new partition is given by $\tilde{T} = \left(\tilde{T}_1, \tilde{T}_2, \ldots, \tilde{T}_{m-k+1}\right)$, where $\tilde{T}_1 = \{j_1, \ldots, j_k\}$ and $\tilde{T}_i = T_{i+k-1}$ for $i = 2, \ldots, m-k+1$. Then $O^{BC}_{j_l}(W, T) = O^{BC}_{\tilde{T}_1}\left(W, \tilde{T}\right)$ for any $l = 1, \ldots, k$.

This equality follows from the observation that when we compute the value of the coalitional BC index in both case swings of players in singletons are the same as swings of players in the new party $\tilde{T}_1$. In the first case the number of swings is divided by $2^{m+1-2} = 2^{m-1}$ because there are m parties and the cardinality of the singleton is 1. In the second case we divide the number of swings by $2^{m-k+1+k-2} = 2^{m-1}$ since the number of parties is equal to $m-k+1$ and the cardinality of the new party $\tilde{T}_1$ is equal to k.

It means that in our case the value of the coalitional BC index of a voter from the party T_1 depends actually only on the size of the party T_2 and does not depend on the size of the own party (in other words when calculating the power of the first party members using the coalitional BC index we consider the situation where there is only one party of the size t_2 and all remaining voters form singletons). The behavior of this index is shown at Fig. 14.

The coalitional BC index of a member of the first party is then the decreasing function of t_2 and it achieves its maximal value in the situation where $t_2 = 2$ (and t_1 is arbitrary). The shape of this curve is rather similar to the shape of curves in case of composite game (and different from the shape of curves illustrating the behavior of coalitional SS index).

If we want to examine the behavior of both indices regarding their dependence on the size of the own party (with t_2 fixed), then the picture is as it can be seen at Figs. 15 and 16.

In case of coalitional SS index we observe that the power of the member of the first party is (almost) monotonic function of the own party's size. For large sizes of the opponent we notice a slight decrease of the power of a member of T_1, but then the power increases monotonically and achieves maximum always for $t_1 = 50$.

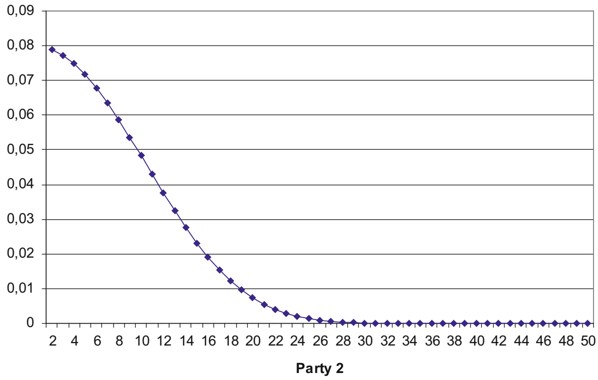


Fig. 14 Coalitional BC index of a member of T_1 (as a function of t_2)

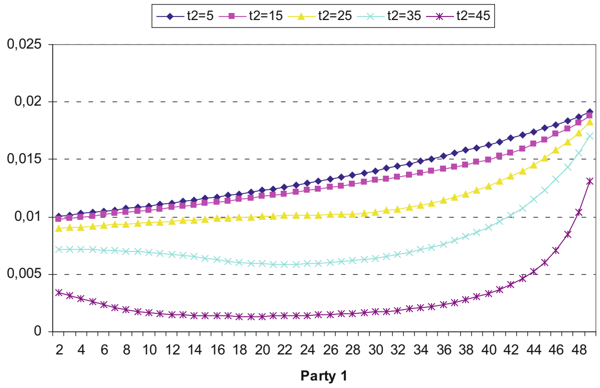


Fig. 15 Coalitional SS index of a member of T_1 (as a function of t_1) for various configurations of sizes of both parties

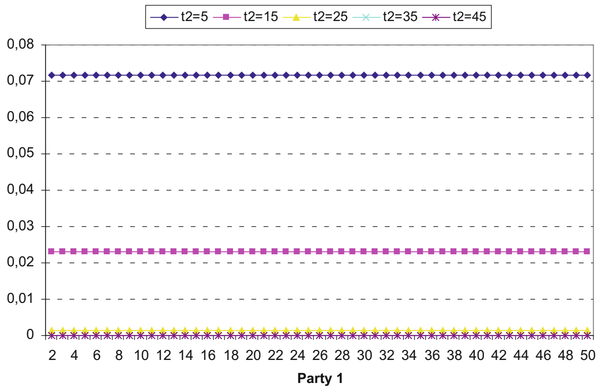


Fig. 16 Coalitional BC index of a member of T_1 (as a function of t_1) for various configurations of sizes of both parties

Besides, we observe that as the opponent party's size increases, corresponding curves are coming down. In case of the coalitional BC index we have horizontal lines, since the power of a voter does not depend on the own party's size, but also lines corresponding to smaller values of t_2 are placed higher.

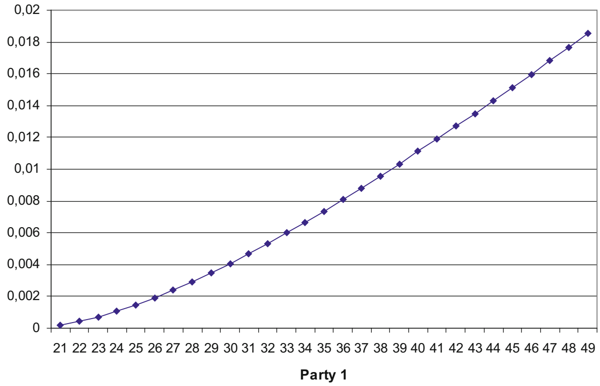


Fig. 17 Coalitional SS index of the member of T_1 (as a function of t_1)—the migration from T_2 to T_1 ($l = 30$)

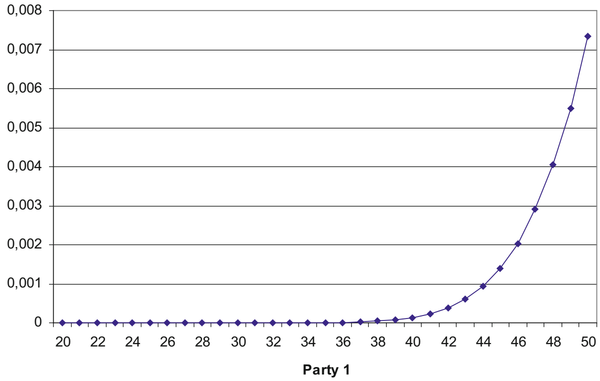


Fig. 18 Coalitional BC index of the member of T_1 (as a function of t_1)—the migration from T_2 to T_1 ($l = 30$)

When considering the migration from the party T_2 to the party T_1 we have the monotonicity result: the larger is the own party's size (and in the mean time the smaller is the size of the opponent), the greater is the power of the first party member measured by both BC and SS coalitional indices. Figures 17, 18, 19, and 20 show this result.

And finally we come to the results concerning individual voters, which are similar to the case of composite game. The power of individual voter, measured by coalitional SS index is a symmetric function of the size of one party (keeping the number of individual voters constant) and attains its maximum in the situation where both parties are of the same size (or their sizes differ by 1 member). An example of the behavior of the power of an individual voter measured by coalitional SS index is shown at Fig. 21. We do not consider here the coalitional BC index of an individual voter since we argued that it is the same that his BC index in the composite game.

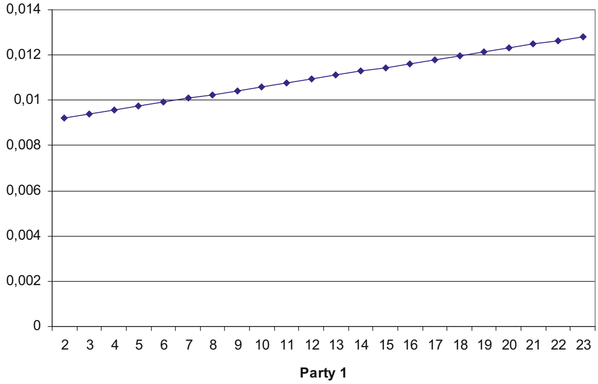


Fig. 19 Coalitional SS index of the member of T_1 (as a function of t_1)—the migration from T_2 to T_1 ($l = 75$)

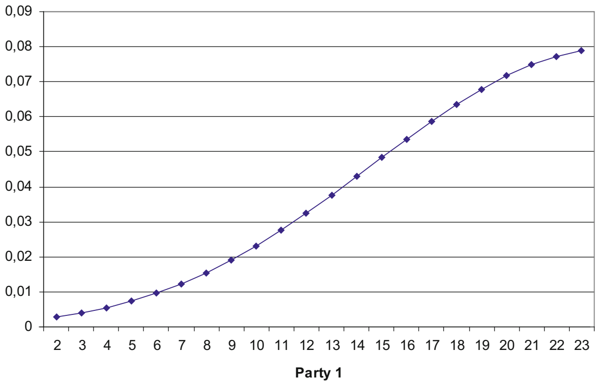


Fig. 20 Coalitional BC index of the member of T_1 (as a function of t_1)—the migration from T_2 to T_1 ($l = 75$)

6 Comparison of the Same Indices in Different Games

We thought that it could also be interesting to compare the behavior of the SS index and BC index of a voter in two different games (which means using two alternative ways of measuring the power of a voter in a voting body divided into parties). First we show some results concerning the SS index. We compare the range and shape of curves corresponding to the power of party members treated as a function of the size of the opponent (Figs. 22 and 23).

We notice that considered indices behave in different way. The range of the coalitional SS index is less than the range of the SS index in a composite game. Coalitional SS index is almost constant for small sizes of the opponent party and then it decreases rather slowly. SS index in a composite game is considerably greater

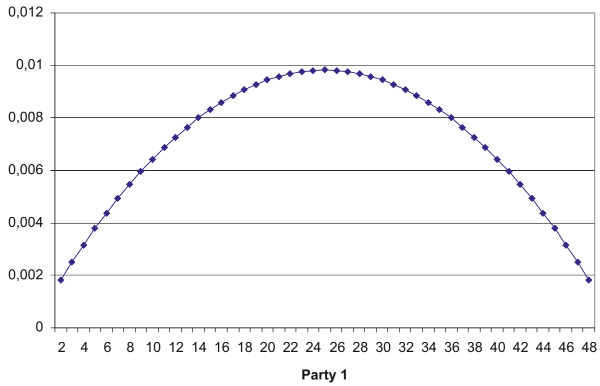


Fig. 21 Coalitional SS index of an independent voter as a function of the size T_1 for $l = 50$

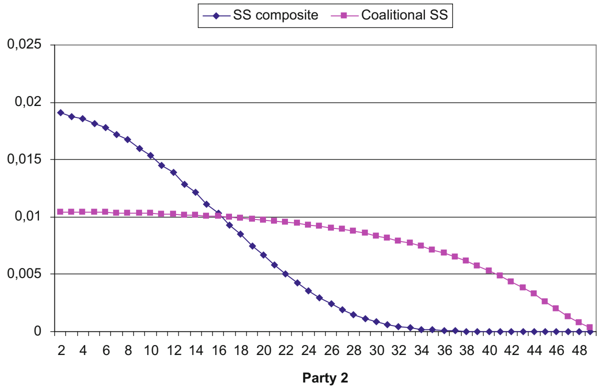


Fig. 22 The power (SS) of the first party's member as a function of t_2 with $t_1 = 5$

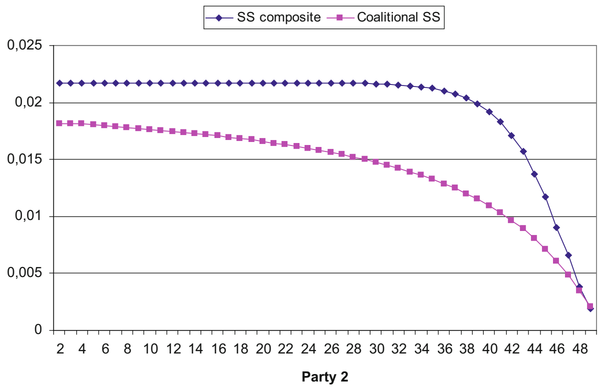


Fig. 23 The power (SS) of the first party's member as a function of t_2 with $t_1 = 46$

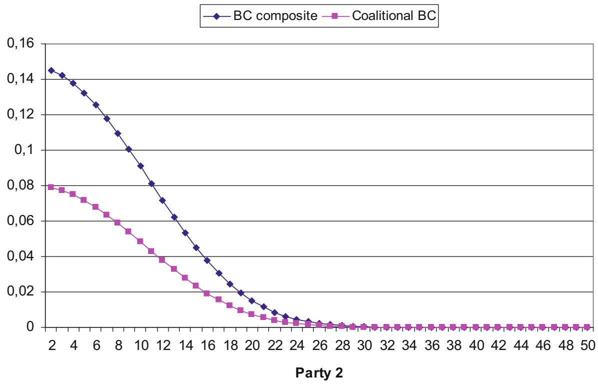


Fig. 24 The power (BC) of the first party's member as a function of t_2 with $t_1 = 5$

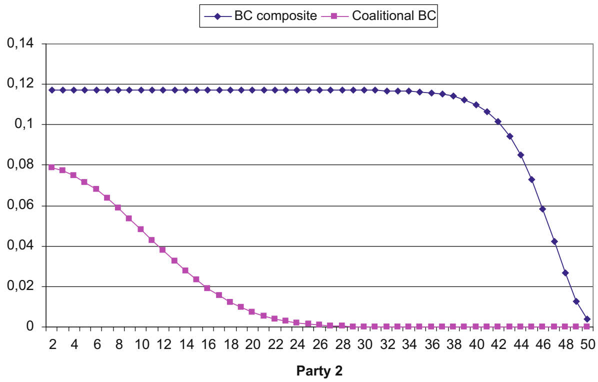


Fig. 25 The power (BC) of the first party's member as a function of t_2 with $t_1 = 46$

than coalitional SS index for small sizes of the opponent party (for t_1 greater than 46 SS index in a composite game is greater than coalitional SS index for all possible sizes of the second party). For small t_1 the index Sh^c decreases quickly with the increase of the size of the second party, it achieves the level of the O^{SS} index, then it has an inflection point and it decreases slowly to the values close to zero. For larger t_1 the behavior of the Sh^c index is different. For small sizes of the opponent party it is almost constant and starts to decrease as the size of the second party is quite large. The point of intersection with the O^{SS} index curve moves to the right (to the larger sizes of the second party) with the increasing size of the first party and eventually Sh^c index is larger than O^{SS} index for all possible values of t_2 .

What is the picture if we compare the behavior of BC index in two different frameworks? It turns out that conclusions are different (Figs. 24 and 25).

Again the range of the β^c index is greater than the range of the O^{BC} index, but here the value of β^c index is (almost) always greater than the value of the O^{BC} index (the equality occurs in case where $t_1 = 2$).

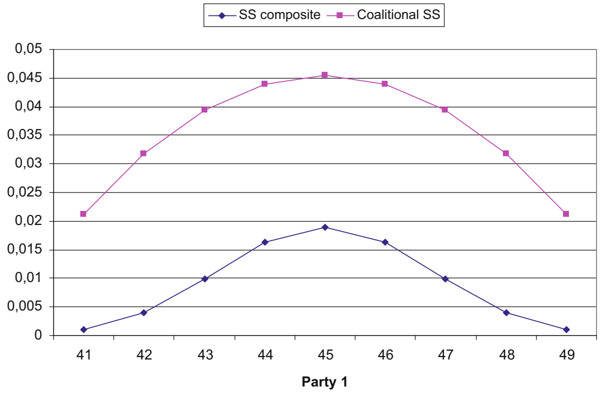


Fig. 26 The power (SS) of an independent voter as a function of the size T_1 for $l = 10$

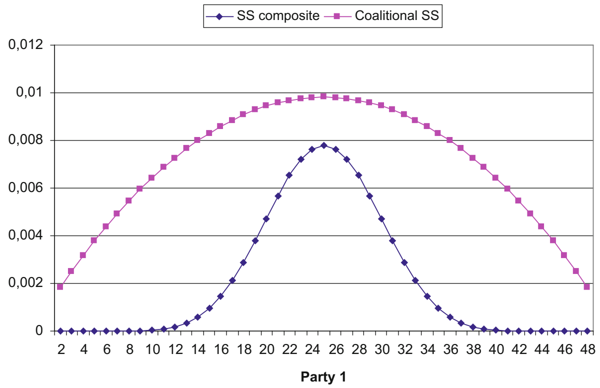


Fig. 27 The power (SS) of an independent voter as a function of the size of T_1 for $l = 50$

If we measure the power of an individual voter by means of the SS index in two different approaches (composite game versus game with a coalition structure), then in turns out that the power of an individual voter is always greater in game with precoalitions than in the composite game (Figs. 26 and 27).

In case of BC index we do not have such conclusion, because the value of BC index of an individual voter is the same in both games. What we can conclude is that the relative BC power if an independent voter in a composite game is less than in the game with precoalitions since the BC power of party members in the composite game is greater than in game with precoalitions.

Note that if we do not consider coalition structure in the voting body, then all voters have the same voting power (considering any majority voting rule and any symmetric index). In case of SS index the power of each individual voter in the concerned voting body is equal to 0.01. We can ask the following questions: when the party membership increases the power of a voter or for which coalition structures

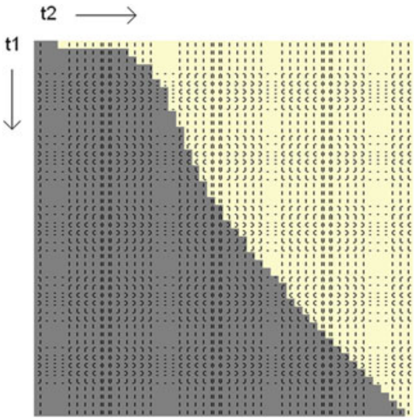


Fig. 28 SS index in a composite game of a member of T_1—comparison with a non-party case

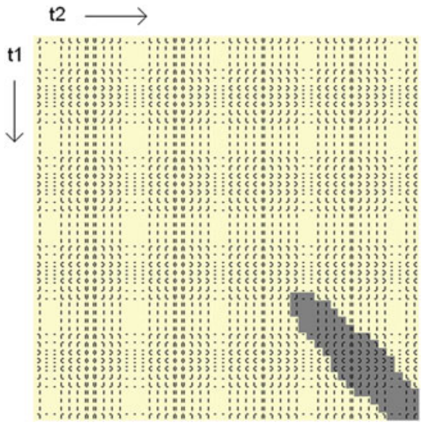


Fig. 29 SS index in a composite game of an independent voter—comparison with a non-party case

the power of an independent voter is greater than in the situation when the coalition structure does not exists. The answer to those questions for both indices–Sh^c and O^{SS}–is given in Figs. 28, 29, 30, and 31. At each figure there are shown values of respective index for a member of T_1 or for an independent voter at all possible configurations of sizes of both parties (rows correspond to the size of the first party, and columns correspond to the size of the second party; the left upper corner corresponds to the case $t_1 = t_2 = 2$ while the right bottom corner describes the case $t_1 = t_2 = 50$). The cells are shaded if the value of respective index is greater than 0.01.

Looking at those pictures we conclude that for both indices taking into consideration the coalition structure in most cases increases the power of a party member (comparing to the case without any a priori coalition structure). If we take the index Sh^c, then for small t_1 the power of a party member is less than 0.01 in case where t_1 is substantially less than t_2. For larger values of t_1 the power of T_1's member becomes less than 0.01 if the size of the second party is greater than the size the first one. Independent voters are better off when considering the party structure

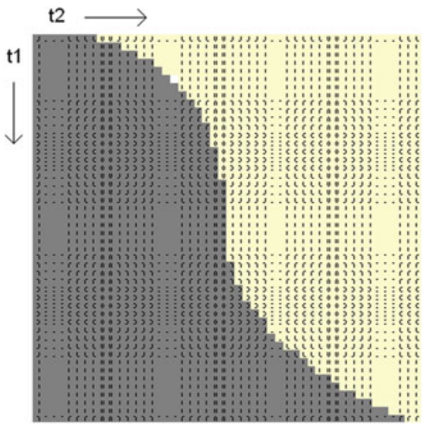


Fig. 30 Coalitional SS index of a member of T_1—comparison with a non-party case

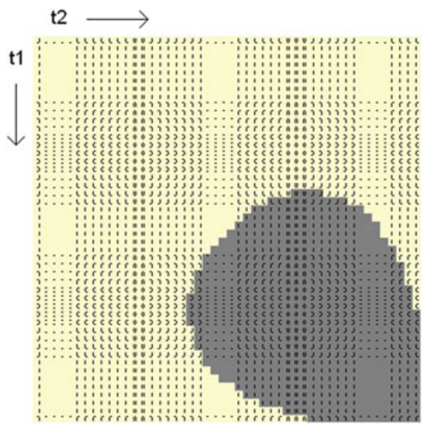


Fig. 31 Coalitional SS index of an independent voter—comparison with a non-party case

only in cases where both parties are approximately of the same size and both are rather large.

The last observation is such that the coalitional SS index in general promotes independent voters while the SS index in a composite game gives more power to the party members.

Similar analysis for the BC index leads to a bit different conclusions (now we compare the value of BC index in considered games with the value of BC index of a voter in a 100-person simple majority voting game). In case of composite games the situation is analogous, but in case of games with a priori unions party members are always worse off comparing to the case of lack of parties, while the situation of independent voters does not change comparing to composite games. Figures 32, 33, and 34 illustrating this issue are given below.

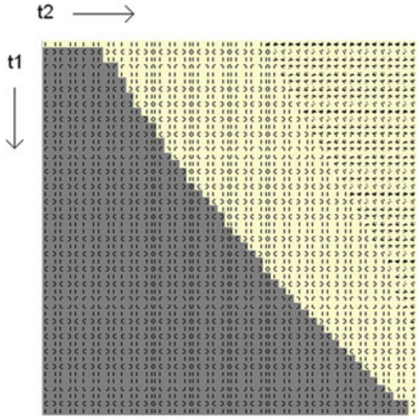


Fig. 32 BC index of a member of T_1 in a composite game—comparison with a non-party case

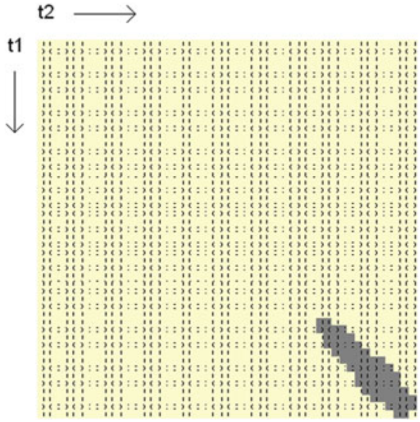


Fig. 33 BC index of an independent voter in a composite game (and in a game with precoalitions)—comparison with a non-party case

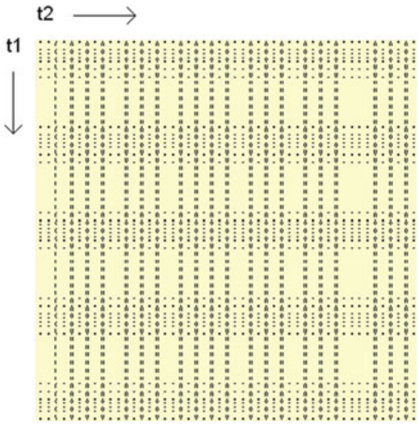


Fig. 34 Coalitional BC index of a member of T_1—comparison with a non-party case

7 Concluding Remarks

The comparison of both indices in the considered case of a voting body leads to some conclusions concerning the properties of both methods of measuring the voting power of individuals in a voting body with a coalition structure. First of all we observe that both indices in a composite game are more sensitive to the changes of coalition structure and have larger range of values than their counterparts in a game with the coalition structure. On the other hand, in games with precoalitions Owen modifications of SS index and BC index are in most cases monotone with respect to the size of ones own party and the size of the opponent. If the size of the opponent is arbitrarily fixed, then maximal power is always achieved while own party's size is maximal (= 50); if ones own party has an arbitrarily settled size, then the power (measured by both indices) of its member is a decreasing function of the opponent's size. Moreover, the larger is own party's size, the larger is maximal possible power of its member, which means, that the global maximum is attained while the own party has 50 members and the opponent has two members. In a composite game indices Sh^c and β^c do not reveal such monotonicity. While the own party's size is fixed, the power of its member is also a decreasing function of the opponent's size. However, with the arbitrarily fixed size of the opponent party, the maximum of power depends on the size of the opponent. The global maximum is achieved in the situation where the own party has 22 members and the opponent has 2 members in case of the Sh^c index and for $t_1 = 13$ and $t_2 = 2$ in case of the β^c index.

Independent voters are more powerful in game with precoalitions than in a composite game.

The conclusion which raises after the analysis of our simulations is that the behavior of SS index and BC index depends much more on the structure of the game considered than on the index itself which implies the fact that the behavior of SS index in a composite game is much more similar to the behavior of the BC index in that game than to the behavior coalitional SS index in a game with precoalitions.

Another issue is the interpretation and, in consequence, the choice of one of described here measures of power (and a proper model) for applications. A criterion which could be helpful is the discipline of voting in parties. If there is a party whip, then the model of composite game should be applied (especially in case of BC index). Notice that the obvious interpretation of the index β^c is that we deal with a situation where all members of each party follow the discipline and vote according to the decision made by internal voting. In case of the β^c index the power of a voter decomposes into two factors–one is the individual power in the internal voting and second is the power of a party as a whole. Relations between these two factors were examined in Leech and Leech (2006). On the other hand we can interpret the index O^{BC} as a measure of power of a member of a party where there is no party whip, assuming that in all other parties voters follow the party discipline. We obtained in our simulations a result that it is always better for members of the disciplined party when their opponents do not have a party whip. It is worth noting at this point that the choice of the voting model may depend on the subject voted, because discipline of voting inside a given party is usually demanded or not subject to the topic under consideration.

From the numerical point of view the calculation of the BC index is much simpler than of SS index especially in case of composite game. On the other hand the fact that the coalitional SS index can be decomposed into two factors provides quite easy way of obtaining numerical results. We restricted our research to the case of two parties because it allowed for an illustrative presentation of results. Obviously, the methodology presented here can be applied to examine the power of members of actual voting bodies with an arbitrary structure of parties.

Acknowledgements I would like to express my gratitude to the reviewer of this paper for helpful comments and suggestions of amendments.

References

Albizuri, M. J. (2000). *An axiomatization of the modified Banzhaf-Coleman index*. Discussion Papers, Department of Applied Economics IV Basque Country University, No 8.

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

Coleman, J. S. (1964). *Introduction to mathematical sociology*. New York: Free Press of Glencoe.

Dubey, P. (1975). On the uniqueness of the Shapley value. *International Journal of Game Theory, 4*, 131–139.

Ekes, M. (2006). Two types of the Banzhaf–Coleman index in games with a priori unions. *Roczniki Kolegium Analiz Ekonomicznych SGH, zeszyt, 15*, 31–45.

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

Hart, S., & Kurz, M. (1983). On the endogenous formation of coalitions. *Econometrica, 51*, 1295–1313.

Laruelle, A., & Valenciano, F. (2004). On the meaning of Owen-Banzhaf coalitional value in voting situations. *Theory and Decision, 56*, 113–123.

Leech, D., & Leech, R. (2006). Voting power and voting blocs. *Public Choice, 127*, 285–303.

Owen, G. (1977). Values of games with a priori unions. In R. Henn & O. Moschlin (Eds.), *Lecture notes in economics and mathematical systems. Essays in honour of Oskar Morgenstern* (pp. 76–88). New York: Springer.

Owen, G. (1978). Characterization of the Banzhaf-Coleman index. *SIAM Journal of Applied Mathematics, 35*, 315–327.

Owen, G. (1981). Modification of the Banzhaf-Coleman index for games with a priori unions. In M. J. Holler (Ed.), *Power, voting and voting power* (pp. 232–238). Wurzburg: Physica.

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

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

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

Pathology or Revelation? The Public Good Index

Manfred J. Holler and Hannu Nurmi

1 The PGI Introduction

This paper focuses on the representation of causality in collective decision making by means of power and power measures and discusses the question whether the Public Good Index (PGI) is a suitable instrument for this representation. To answer this question we relate the PGI and, alternatively, the Banzhaf index to the NESS concept of causality. However, it shows that the answer also depends on whether we interpret the PGI as measure or as an indicator.

Section 2 discusses the well-known fact that the PGI violates the axiom of local monotonicity (LM), i.e., it is not guaranteed that a player that can make a larger contribution to winning than another has at least as large an index value. In Sect. 3, we argue that cases of nonmonotonicity indicate properties of the underlying decision situations which cannot be brought to light by the more popular power measures, i.e., the Banzhaf index and the Shapley–Shubik index, that satisfy LM. The discussion proposes that we can constrain the set of games representing decision situations such that LM also holds for the PGI. This might be a helpful instrument for the design of voting bodies. The discussion of causality in Sect. 4 suggests that the nonmonotonicity can be the result of framing the decision problem in a particular way and can perhaps even ask the "wrong question". The core of this section is dedicated to connecting power and responsibility in the case of collective decision making and collective action, i.e., when the cause for an outcome cannot

M.J. Holler (✉)
Department of SocioEconomics, University of Hamburg, Hamburg, Germany

Public Choice Research Centre, University of Turku, Turku, Finland
e-mail: holler@econ.uni-hamburg.de

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

Public Choice Research Centre, University of Turku, Turku, Finland

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare, DOI 10.1007/978-3-319-05158-1_14,

be directly assigned to a particular individual agent. Based on the discussion in the previous sections, Sect. 5 points out that the PGI can be interpreted as an indicator and thus even serve as a valuable instrument in cases where there are serious doubts raised whether it can be applied as a measure. To conclude, Sect. 6 looks into the probabilistic relationship of Banzhaf index and PGI as elaborated independently by Widgrén (2002) and Brueckner (2002) that identifies the factor which is responsible for the formal difference between the two measures. Can we interpret this factor as the cause for the violation of LM that characterizes the PGI, but not the Banzhaf? Can we see from the properties of this factor whether the PGI will indicate a violation for a particular game, or not? However, these are questions that have not been answered as yet.

The normalized Banzhaf index of player i counts the number of coalitions S that have i as a swing player such that S is a winning coalition and $S \setminus \{i\}$ is a losing coalition for all $S \subset N$ if N is the set of all players of game v. For normalization this number is divided by the total number of swing positions that characterize the game v.

The PGI differs from the Banzhaf index inasmuch as only minimum winning coalitions (MWCs) are considered. S is a MWC if $S \setminus \{i\}$ is a losing one, for all i (S, i.e., all players of a MWC have a swing position. The PGI of player i, h_i, counts the number of MWCs that have i as a member and divides this sum by the sum of all swing positions the players have in all MWCs of the game. If m_i is the number of MWCs that have i as a member then is PGI value is

$$h_i = \frac{m_i}{\sum_{i \in N} m_i} \tag{1}$$

The corresponding definition of the normalized Banzhaf index is

$$\beta_i = \frac{c_i}{\sum_{i \in N} c_i} \tag{2}$$

In (2), c_i is number of winning coalitions that have i as a swing player. The following analysis is based on these two power measures.

2 The Pathology

In 1978, when Holler first applied the PGI to the study of a voting power distribution in a parliament, he concluded that facing the violation of LM "causes doubt" concerning the validity of this measure. Obviously, he found the "index of Banzhaf–Coleman type" which he used as an alternative "more adequate in the context of this analysis" (Holler 1978, p. 33). However, Holler (1982) argued that taking into consideration coalition formation, collective decision making and the public goods problem the focus on MWCs and thus on the PGI seems to be an adequate solution to measuring the distribution of voting power in the decision making body. This view was supported by the axiomatization of the PGI in Holler and Packel (1983) and Napel (1999). However, in their article "Postulates and Paradoxes of Relative Voting Power—A Critical Review", Dan Felsenthal and Moshé Machover (1995, p. 211) write that "it seems intuitively obvious that if $w_i \leq w_j$ then every voter j has at least as much voting power as voter i, because any contribution that i can make to the passage of a resolution can be equalled or bettered by j." They conclude that "any reasonable power index" should be required to satisfy local monotonicity, i.e., LM. Even more distinctly, they argue that any a priori measure of power that violates LM is "pathological" and should be disqualified as a valid yardstick for measuring power (Felsenthal and Machover 1998, p. 221ff). This argument has been repeated again and again when it comes to the evaluation (and application) of the PGI and the Deegan–Packel index.[1]

A notorious example to illustrate the nonmonotonicity of the PGI is the voting game $v^0 = (51; 35, 20, 15, 15, 15)$. The corresponding PGI is

$$h(v^0) = (\frac{4}{15}, \frac{2}{15}, \frac{3}{15}, \frac{3}{15}, \frac{3}{15}),$$

indicating a violation of LM in the resulting distribution of a priori voting power.[2]

The application of power indices is motivated by the widely shared "hypothesis" that the vote distribution is a poor proxy for a prior voting power. If this is the case, does it make sense to evaluate a power measure by means of a property that refers to the vote distribution as suggested by LM? Of course, our intuition supports LM. However, if we could trust our intuition, do we need the highly sophisticated power measures at all?[3]

3 The Revelation

It has been argued that a larger voter j can be less welcome to join a (non-winning) proto-coalition than a smaller voter i.[4] The intuitive argument is the following. Let's assume a voting game $v* = (51; 45, 20, 20, 15)$ and players 2 and 3 form a proto-coalition $S = \{2, 3\}$. The losing coalition S can be "transformed" into a winning

[1]The Deegan–Packel index was introduced in Deegan and Packel (1979). This measure considers the value of coalition to be a private good that is equally shared among the members of a coalition. For a recent discussion of this measure, taking a priori unions into account, see Alonso-Meijide et al. (2010, 2011) and Holler and Nohn (2009).

[2]The corresponding Deegan–Packel index, $\rho(v^0) = (18/60, 9/60, 11/60, 11/60, 11/60)$, also shows a violation of LM.

[3]See Holler (1997) and Holler and Nurmi (2010) for this argument.

[4]For a discussion of coalition formation, see e.g. Hardin (1976), Hart and Kurz (1984), Holler and Widgren (1999), Holler (2011), Miller (1984), and Riker (1962).

coalition if either player 1 or player 4 or both join S. However, if player 1 joins, either individually or together with 4, then neither player 2 nor player 3 is critical to the winning of a majority, i.e., in the coalition $\{1, 2, 3\}$ neither 2 nor 3 is a swinger. If voting power refers to a swing position—and this is, with some modification, the kernel of all standard power measures—and players are interested in power, then it seems likely that players 2 and 3 prefer the "smaller" voter 4 to join S to form a winning coalition. This story tells us that it could well be that a larger player is not always welcome to form a winning coalition if a smaller one does the same job. But does this mean that only minimum winning coalitions will form? Empirical evidence speaks against this conclusion. However, it has been repeatedly argued that if (nonminimal) winning coalitions with surplus players form then this is due to luck or ideology (i.e. preferences) and should not be taken into consideration when it comes to represent a priori voting power.[5] But there are perhaps more straightforward arguments in favor of MWCs and the application of the PGI.

In Holler and Napel (2004a,b) it has been argued that the PGI shows nonmonotonicity with respect to the vote distribution (and thus confirms that the measure does not satisfy LM) if the game is not decisive, as the above weighted voting game $v^0 = (51; 35, 20, 15, 15, 15)$, or improper (for an example, see Sect. 4 below) and therefore indicates that perhaps we should worry about the design of the decision situation. The more popular power measures, i.e., the Shapley–Shubik index or the Banzhaf index satisfy LM and thus do not indicate any particularity if the game is neither decisive nor proper. Interestingly, these measures also show a violation of LM if we consider a priori unions and the equal probability of permutations and coalitions, respectively, does no longer apply.[6] This suggests that a deviation of the equal probability of coalitions causes a violation of LM.

The concept of a priori unions or pre-coalition is rather crude when applied to the PGI as the PGI implies that certain coalitions will not be taken into consideration at all, i.e., have a probability of zero of forming. Note since the PGI considers MWC only, this is formally equivalent to put a zero weight on coalitions that have surplus players. Is this the ("technical") reason why the PGI may show nonmonotonicity? We will come back to this question in Sect. 6 below.

Instead of accepting the violation of LM, we may ask which decision situations guarantee monotonic results for the PGI. An answer to this question may help to design adequate voting bodies. Obviously, the PGI satisfies LM for unanimity games, dictator games and symmetric games. The latter are games that give equal power to each voter; in fact, unanimity games are a subset of symmetric games. Note that for these types of games the PGI is identical with the normalized Banzhaf index.

---

[5] See Holler (1982) and Holler and Packel (1983). See also Widgrén (2002).

[6] See Alonso-Meijide and Bowles (2005) for examples of voting games with a priori unions and Alonso-Meijide and Holler (2009), Alonso-Meijide et al. (2009) as well as Holler and Nurmi (2010) for a discussion. See also Alonso-Meijide et al. (2010) and Holler and Nohn (2009).

In Holler et al. (2001), the authors analyze alternative constraints on the number of players and other properties of the decision situations. For example, it is obvious that local monotonicity will not be violated by any of the known power measures, including PGI, if there are n voters and $n - 2$ voters are dummies. It is, however, less obvious that local monotonicity is also satisfied for the PGI if one constrains the set of games so that there are only $n - 4$ dummies. A hypothesis that needs further research is that the PGI does not show nonmonotonicity if the voting game is decisive and proper and the number of decision makers is smaller than 6. (Perhaps this result also holds for a larger number of decision makers but we do not know of any proof.) The idea of restricting the set of games such that LM applies for PGI has been further elaborated in Alonso-Meijide and Holler (2009) in the form of "weighted monotonicity of power." It seems that these considerations are relevant for all power indices if we drop the equal probability assumption and, for example, take the possibility of a priori unions into account.

4 Causality and Power

The elaboration of various power measures and their discussion is meant to increase our understanding of power in collectivities and also to be of help in the design of voting bodies. A relatively new application of these measures results from their formal equivalence with representations of causality in collective decision making. Given this, it seems a short step to equate power and responsibility.

The specification of causality in the case of collective decision making with respect to individual agents cannot be derived from the action and the result as both are determined by the collectivity. They have to be traced back to decision making and, in general, the decision making process. However, collective decision making has a quality that substantially differs from individual decision making. For instance, an agent may support his favored alternative by voting for another alternative or by not voting at all. Nurmi (1999, 2006) contain a collection of such "paradoxes". These paradoxes tell us that we cannot derive the contribution of an individual to a particular collective action from the individual's voting behavior.[7] Trivially, a vote is not a contribution, but a decision. Resources such as power, money, etc. are potential contributions and causality might be traced back to them if a collective action results. As a consequence causality follows even from votes that do not support the collective action. This is reflected by everyday language that simply states that the Parliament has decided when in fact a decision was made by a majority smaller than 100 %. But how can we allocate causality if it is not derived from decisions?

Alternatively, we may assume in what follows that the vote (even in committees) is secret and we do not know who voted "yes" or "no". Moreover, in general, there

[7]For a discussion and examples, see Nurmi (2010) and Holler and Nurmi (2014).

are more than two alternatives and the fact that a voter votes "yes" for A in a last pairwise voting only means that he/she prefers A to B or does not want to abstain, but this vote does not tell us why and how alternatives C, D, etc. were excluded.[8] Thus, an adequate concept of causality (and responsibility) does not presuppose a voting result that is known and indicates who said "yes" and who said "no".[9]

For an illustration, imagine a five-person committee $N = \{1, 2, 3, 4, 5\}$ that makes a choice between the two alternatives x and y.[10] The voting rule specifies that x is chosen if either (a) 1 votes for x, or (b) at least three of the players 2–5 vote for x. Let's assume all individuals vote for x. What can be said about causality? Clearly, this is a case of over-determination and the allocation of causation is not straightforward. Alternatively, we may assume that all we get to know is that x is decided, but we do not know who voted for or against it. In both cases, we may conclude on causality by looking at possible winning coalitions. For example, the action of agent 1 is a member of only one minimally sufficient coalition, i.e. decisive set, while the actions of each of the other four members are in three decisive sets each. If we take the membership in decisive sets as a proxy for causation, and standardize such that the shares of causation adds up to one, then vector

$$h^0 = (\frac{1}{13}, \frac{3}{13}, \frac{3}{13}, \frac{3}{13}, \frac{3}{13})$$

represents the degrees of causation. Braham and van Hees (2009, p. 334), who introduced and discussed the above case, conclude that "this is a questionable allocation of causality." They add that "by focusing on minimally sufficient conditions, the measure ignores the fact that anything that players 2–5 can do to achieve x, player 1 can do, and in fact more—he can do it alone." We share this specification, but does it apply in collective decision making?

Let's review the above example. Imagine that x stands for polluting a lake. Now the lake is polluted, and all five members of N are under suspicion of having contributed to its pollution. Then h^0 implies that the share of causation for 1 is significantly smaller than the shares of causation of each of the other four members of N. If responsibility and perhaps even punishment follow from causation then the allocation h^0 seems highly pathological. As a consequence Braham and van Hees propose to apply the weak NESS instead of the strong one, i.e., not to refer to decisive sets, but to consider sufficient sets instead and count how often an element i of N is a necessary element of a sufficient set (i.e., a NESS). Taking care of an adequate standardization so that the shares add up to 1, we get the following allocation of causation:

[8]See "The Fatal Vote: Berlin versus Bonn" (Leininger 1993) for an illustration.

[9]See Braham (2005, 2008), Braham and Holler (2009), and Holler (2007) for this treatment of causality. It differs from the approach discussed in Felsenthal and Machover (2009) which refers to particular (voting) *results*, and not for the *potential* of having contributed to it.

[10]The rest of this section derives from Holler (2012).

$$\beta^0 = (\frac{11}{23}, \frac{3}{23}, \frac{3}{23}, \frac{3}{23}, \frac{3}{23}).$$

The result expressed by b^0 looks much more convincing than the result proposed by h^0, doesn't it? Note that the b-measure and h-measure correspond to the Banzhaf index and the PGI, respectively, and can be calculated accordingly.

So far the numerical results propose the weak NESS test and thus the application of the normalized Banzhaf index. However, what happened to alternative y? If y represents "no pollution" then the set of decisive sets consists of all subsets of N that are formed of the actions of agent 1 and the actions of two out of agents $2, 3, 4$ and 5. Thus, the actions of 1 are members of six decisive sets while the actions of $2, 3, 4$ and 5 are members of three decisive sets each. The corresponding shares are given by the vector

$$h* = (\frac{1}{3}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6})$$

Obviously, $h*$ looks much more convincing than h^0 and the critical interpretation of Braham and van Hees does no longer apply: agent 1 cannot bring about y on its own, but can cooperate with six different pairs of two other agents to achieve this goal.

Note that the actions (votes) bringing about x represent an improper game—two (disjunct) "winning" subsets can exist at the same time[11]—while the determination of y can be described by a proper game. However, if there are only two alternatives x and y then "not x" necessarily implies y, irrespective of whether the (social) result is determined by voting or by polluting. The h-values indicate that it seems to matter what issue we analyze and what questions we raise while the Banzhaf index with respect to y is identical to the one for $x : \beta^0 = b*$.

Already Coleman (1971) developed a measure of power of a member to initiate action and a measure of power of a member to prevent action. However, Brams and Affuso (1976) demonstrated that, submitted to adequate linear transformations, the two measures yield the normalized Banzhaf index. This is not the case with public good values h^0 and $h*$ above.

5 Measure or Indicator?

Whether we should apply h and β, or a third alternative, to measure causation seems an open question, and this paper will not give an answer to this question. To conclude, the PGI and thus the strong NESS concept may produce results that are counter-intuitive at first glance. However, in some decision situations they seem

[11] Player 1 is a dictator in guaranteeing x, but x can also be achieved without his support.

to reveal more about the power structure and corresponding causality allocation than the Banzhaf index and the corresponding weak NESS concept. However, if we want to relate responsibility to power then the nonmonotonicity, i.e. the violation of LM, that represents the strong NESS test of the PGI is quite a challenge: If the collective choice is made through voting then it is not guaranteed that a voter with a larger share of votes has at least as much responsibility for the collectively determined outcome as a voter with a smaller share. From the example above we can learn that nonmonotonicity might indicate that we asked the wrong question: Is the responsibility with respect to keeping the lake clean or is it with polluting the lake? Both alternatives may imply the sharing of the costs of cleaning it. Of course, there is no quantitative answer to this question, but the quantification by the index showed us that there might be a problem with the specification of the game model.

A possible answer of whether the PGI represents a pathology or not, might be found in this quality–quantity duality: the use of quantity measures to indicate qualitative properties of (voting) games. Whether a game is improper or non-decisive is not a matter of degree. Indicators show red lights or make strange noises when an event happens that has some meaning in a particular context. This does not necessarily mean that the corresponding indicator functions as a measure, but often it does and when it does it summarizes the measured values in the form of signals. What is a relevant and an appropriate signal of course depends on the context and the recipient. Red lights are not very helpful for blind people. What are the relevant and appropriate signals that correspond to power measures? What are the problems that should be uncovered and perhaps even be solved? What are the properties a power measure has to satisfy when it should serve as a signal? These are questions that we cannot answer in a systematic way without reference to a particular issue.

6 On the Relationship of Banzhaf Index and PGI

Widgrén (2002) proved the following linear relationship that relates the normalized Banzhaf index (β_i) and the PGI (h_i).[12]

$$\beta_i = (1-\pi)h_i + \pi\varepsilon_i \tag{3}$$

where

$$\varepsilon_i = \frac{\overline{c_i}}{\sum_{i \in N} \overline{c_i}} \text{ and } \pi = \frac{\sum_{i \in N} \overline{c_i}}{\sum_{i \in N} c_i}$$

[12]Widgrén uses the symbols θ_i for the PGI and C_i for the set of crucial coalitions that contain i as a swing player. Correspondingly, c_i is the number of elements of C_i.

Here, c_i represents the number of (crucial) coalitions that contain player i as a swing player and $\overline{c}_i$ represents the number of coalitions which have a swing player i, but are *not* minimum winning. If we apply Eq. (3) to the voting game x discussed in Sect. 4 and the corresponding power indices β^0 and h^0, then we have $c_1 = 11, c_2 = c_3 = c_4 = c_5 = 3, \bar{c_1} = 10$ and $\bar{c_i} = 0$ for $i = 2, 3, 4, 5$. As a consequence, $\pi = \frac{10}{23}, \varepsilon_1 = \frac{10}{10}$, and $\varepsilon_i = 0, i = 2, 3, 4, 5$. It is easy to check that these values are consistent with Eq. (3) and the values of β^0 and h^0.

Loosely speaking, the coalitions represented by are the source of the difference between the normalized Banzhaf index, β_i, and the PGI, h_i. Can we identify the corresponding factors in (3) as the cause for the violation of LM that characterizes the PGI, but not the Banzhaf? Can we see from the properties of this factor whether the PGI will indicate a violation for a particular game, or not?—These questions have not been answered so far, but it is immediate from (3) that the PGI satisfies LM for unanimity games, dictator games and symmetric games. For these games $\pi = 0$ and the PGI equals the normalized Banzhaf index (which satisfies LM for all voting games).

Acknowledgements This paper was presented to The Leverhulme Trust sponsored Voting Power in Practice Symposium at the London School of Economics, 20–22 March 2011. The authors would like to thank the participants of the symposium and an anonymous reviewer for helpful comments.

References

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

Alonso-Meijide, J. M., & Holler, M. J. (2009). Freedom of choice and weighted monotonicity of power. *Metroeconomica, 60*(4), 571–583.

Alonso-Meijide, J. M., Bowles, C., Holler, M. J., & Napel, S. (2009). Monotonicity of power in games with a priori unions. *Theory and Decision, 66*, 17–37.

Alonso-Meijide, J. M., Casas-Mendez, B., Fiestas-Janeiro, G., Holler, M. J., & Nohn, A. (2010). Axiomatizations of public good indices with a priori unions. *Social Choice and Welfare, 35*, 517–533.

Alonso-Meijide, J. M., Casas-Mendez, B., Fiestras-Janeiro, G., & Holler, M. J. (2011). The Deegan-Packel index for simple games with a priori unions. *Quality and Quantity, 45*(2), 425–439.

Braham, M. (2005). Causation and the measurement of power. In G. Gambarelli, & M. J. Holler (Eds.), *Power measures III* (*Homo Oeconomicus, 22*) (pp. 645–553). Munich: Accedo.

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

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

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

Brams, S. S., & Affuso, P. J. (1976). Power and size: A new paradox. *Theory and Decision, 7*, 29–56.

Brueckner, M. (2002). Extended probabilistic characterization of power indices. In M. J. Holler, & G. Owen (Eds.), *Power measures* (Vol. 2); *Homo Oeconomicus, 19*, 387–398.
Coleman, J. S. (1971). Control of collectivities and the power of a collectivity to act. In B. Liebermann (Ed.), *Social choice* (pp. 269–301). New York: Gordon and Breach.
Deegan, J., Jr., & Packel, E.W. (1979). A new index of power for simple n-person games. *International Journal of Game Theory, 7*, 113–123.
Felsenthal, D., & Machover, M. (1995). Postulates and paradoxes of relative voting power - a critical appraisal. *Theory and Decision, 38*, 195–229.
Felsenthal, D., & Machover, M. (1998). *The measurement of voting power. Theory and practice, problems and paradoxes*. Cheltenham: Edward Elgar.
Felsenthal, D. S., & Machover, M. (2009). Necessary and sufficient conditions to make the numbers count. *Homo Oeconomicus, 26*, 259–271.
Hardin, R. (1976). Hollow victory: The minimum winning coalition. *American Political Science Review, 70*, 1202–1214.
Hart, S., & Kurz, M. (1984). Stable coalition structures. In: M. J. Holler (Ed.), *Coalitions and collective action*. Wuerzburg: Physica, 235–258.
Holler, M. J. (1978). A prior party power and government formation. *Munich Social Science Review, 4*, 25–41 (Republished in M. J. Holler (Ed.), *Power, voting, and voting power*. Würzburg: Physica, 1982, pp. 273–282.)
Holler, M. J. (1982). Forming coalitions and measuring voting power. *Political Studies, 30*, 262–271.
Holler, M. J. (1997). Power, monotonicity and expectations. *Control and Cybernetics, 26*, 605–607.
Holler, M. J. (2007). Freedom of choice, power, and the responsibility of decision makers. In J.-M. Josselin, & A. Marciano (Eds.), *Democracy, freedom and coercion: A law and economics approach* (pp. 22–45). Cheltenham: Edward Elgar.
Holler, M. J. (2011). Coalition theory. In: K. Dowding (Ed.), *Encyclopedia of power*. London: Sage.
Holler, M. J. (2012). EU decision-making and the allocation of responsibility. In T. Eger, & H.-B. Schäfer (Eds.), *Research handbook on the economics of European Union Law* (pp. 53–81). Cheltenham: Edward Elgar.
Holler, M. J., & Napel, S. (2004a). Local monotonicity of power: Axiom or just a property. *Quality and Quantity, 38*, 637–647.
Holler, M. J., & Napel, S. (2004b). Monotonicity of power and power measures. *Theory and Decision, 56*, 93–111.
Holler, M. J., & Nohn, A. (2009). The public good index with threats in a priori unions. *Essays in honor of Hannu Nurmi (Homo Oeconomicus, 26)*, 393–401.
Holler, M. J., & Nurmi, H. (2010). Measurement of power, probabilities, and alternative models of man. *Quality and Quantity, 44*, 833–847.
Holler, M. J., & Nurmi, H. (2014). Aspects of power overlooked by power indices (in this volume).
Holler, M. J., & Packel, E. W. (1983). Power, luck, and the right index. *Journal of Economics (Zeitschrift für Nationalökonomie) 43*, 21–29.
Holler, M. J., & Widgren, M. (1999). The value of a coalition is power. *Homo Oeconomicus, 15*, 485–501 (Reprinted in *Notizie di Politeia, 59*, 2000, 14–28).
Holler, M. J., Ono, R., & Steffen, F. (2001). Constrained monotonicity and the measurement of power. *Theory and Decision, 50*, 385–397
Leininger, W. (1993). The fatal vote: Berlin versus Bonn. *Finanzarchiv, 50*, 1–20.
Miller, N. R. (1984). Coalition formation and political outcomes: A critical note. In M. J. Holler (Ed.), *Coalitions and collective action* (pp. 259–265). Wuerzburg: Physica.
Napel, S. (1999). The Holler-Packel axiomatization of the public good index completed. *Homo Oeconomicus, 15*, 513–520.
Nurmi, H. (1999). *Voting paradoxes and how to deal with them*. Berlin: Springer.
Nurmi, H. (2006). *Models of political economy*. London: Routledge.

Nurmi, H. (2010). On the relative unimportance of voting weights: Observations on agenda-based voting procedures. In M. Cichocki, & K. Zyczkowski (Eds.), *Institutional design and voting power in the European Union* (pp. 171–180). Farnham, Surrey: Ashgate.

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

Widgrén, M. (2002). On the probabilistic relationship between the public good index and the normalized Banzhaf index. In M. J. Holler, & G. Owen (Eds.), *Power measures* (Vol. 2); *Homo Oeconomicus, 19*, 373–386.

On the Possibility of a Preference-Based Power Index: The Strategic Power Index Revisited

Dieter Schmidtchen and Bernard Steunenberg

1 Introduction

The *strategic power index* (hereafter SPI) is an alternative method for evaluating the distribution of power in policy games, which has been introduced several years ago. Whereas traditional power indices are based on the notion that players need to form some kind of majority or winning coalition, the SPI employs the analytical tools of non-cooperative game theory. Key features of decision-making situations, such as actor preferences, the outcome or policy space, as well as the rules of the decision-making process, are part of a game-theoretical analysis, which forms the basis of the calculation of this index. Since the analysis allows players to act strategically, this index is labeled *strategic*. It reflects the power-related features of the position of political actors in a context as modelled in a non-cooperative game.[1] Since our first proposal of this index, several comments have been raised against the SPI. In this

[1]This index was presented, in non-normalized form, for the first time in 1996 (see Steunenberg et al. 1996). There is another attempt to develop a strategic power index labelled *strict power index* (Napel and Widgren 2002). As with our index, spatial preferences and strategic agenda setting are its main building blocks. However, in the framework of the strict power index, power is defined as the ability of a player "to change the current state of affair" (Napel and Widgren 2002: 4). Following the reasoning of traditional power indices power relates to the ability of being decisive or pivotal.

D. Schmidtchen (✉)
Center for the Study of Law and Economics, University of the Saarland, Building 31,
PO-Box 15 11 50, 66041 Saarbrücken, Germany
e-mail: csle@rz.uni-sb.de

B. Steunenberg
Department of Public Administration, Leiden University, Campus Den Haag (SBS 2),
Schouwburgstraat 2, 2511 VA Den Haag, Leiden, Netherlands
e-mail: b.steunenberg@cdh.leidenuniv.ed

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare,
DOI 10.1007/978-3-319-05158-1_15,

paper we will discuss these comments and clarify our position on voting indices and the SPI.

The SPI is very different from conventional power indices, such as the Banzhaf, Shapley–Shubik and others. These indices take the set of players and the bare decision-making rules as their domain and measure voting power by the extent to which a player in any collective body that makes yes-or-no decisions by vote may turn a losing coalition into a 'winning' coalition for all mathematically possible permutations of players (Shapley–Shubik index) or the relative number of times a player is decisive in a vote (Banzhaf index).[2] As mentioned by Felsenthal and Machover (1998) conventional voting power analyses are either based on cooperative games[3] or are entirely probabilistic measures.[4] What is measured is a priori power. By doing this, conventional power indices do not take account of positive or negative correlations of players preferences.[5] Comments also point to the limited capability of traditional power indices to model players' strategic interaction and a complicated institutional structure typical for real world decision-making (see Garrett and Tsebelis 1997, 1999a, b; Steunenberg et al. 1996, 1997, 1999; Schmidtchen and Steunenberg 2002).

The SPI rests on a notion of power as the ability 'to get what you want'. Important is to distinguish the modelling of a decision-making situation—a game—from the calculation of the index. While the modelling of decision-making situations may be done in different ways—for instance, by introducing new players, changing the sequence of play, information sets, or action sets of players—the calculation

[2]These 'classical' indices have been supplemented with more recent power measures, such as the Johnston index, the Deegan–Packel index and the Holler index. The main differences between these indices are the ways in which coalition members share the benefits of their cooperation, and the kind of coalition players chose to form (see Colomer 1999). For a comparative investigation of traditional power indices see Felsenthal and Machover (1998), Holler and Owen (2001) and Laruelle and Valenciano (2008).

[3]For example, the Shapley–Shubik index measuring what Felsenthal and Machover call P-power, which posits an office-seeking motivation of voting behavior (see Felsenthal and Machover 1998: 171).

[4]See Penrose (1946), Banzhaf (1965), Coleman (1971, 1986), which take a policy-seeking viewpoint focusing on the degree to which a member's vote is able to influence the outcome of a vote. These indices reflect I-power in the sense of Felsenthal and Machover (1998: 36).

[5]That is not to say that traditional power indices are unable to take account of voters' preferences or spatial voting (see Straffin 1994). In probabilistic characterizations of voting power indices each voter i's probability p_i of voting "yes" on a proposal is a random variable. Taking the p_i as an indicator of the acceptability of a proposal to voter i (see Straffin 1994: 1137), homogeneous as well as heterogeneous preferences can be modelled. If each p_i is chosen independently from the uniform distribution on [0,1] we have the Banzhaf index. The independence assumption means that the acceptability of a proposal to voter i is independent of its acceptability to any other voter j (see Straffin 1994: 1137). Note, that $p_i = 1/2$, which means that voter j voting "yes" is similar to flipping a coin. Note further, that the probability characterization of the Banzhaf index is restricted to its non-normalized version. If random variable p is chosen from the uniform distribution on [0,1], and $p_i = p$ for all i (homogeneity assumption), we have the Shapley–Shubik index. Here the acceptability of a proposal is the same to all voters.

of the index remains the same. The index uses, among others, the distances between outcomes and ideal points, which follows a distribution of preference configurations, including status quo points. The resulting expected or average distance for a player is compared with a 'neutral' or dummy player, which helps to differentiate between a player's success and luck. The smaller the expected distance, the more power is attributed to a player compared to the 'neutral' player who does not have any decision-making rights in the game. The index uses a distribution of states of the world, that is, various combinations of preferences and initial policies (status quo points). It levels out the effect of luck, that is, of being close to the equilibrium outcome in a specific game by using numerous different preference configurations and taking averages. The intuition is that the power of a player resides only in the game form or the rules of a game and not in the way a specific game is played (Steunenberg et al. 1996, 1997, 1999; Schmidtchen and Steunenberg 2002).

The SPI gave rise to several comments in the literature. Garrett and Tsebelis (1999b) argue that the SPI—although an improvement compared to conventional indices—nevertheless suffers from a drawback generated by the statistics used in it. Felsenthal and Machover (2001) proved a theorem stating that the SPI is a modified Banzhaf index. Napel and Widgren (2004: 519) give credit to it being the first unified approach to the measurement of decision-making power in that it combines an ex post analysis of well defined games with the ex ante prospect of being successful in the game form underlying these games. However, this first attempt to provide such a framework is considered to be 'problematic'. Napel and Widgren (2004: 524) point to a potential for confounding power and success "that may, but need not, result from it". In an earlier paper, they speak of a confusion of cause and effect (Napel and Widgren 2002: 2). They claim that "(o)nly for particular distribution assumptions ... luck (is) 'leveled out' by taking averages" (Napel and Widgren 2004: 524). The SPI is judged to be "a good measure of expected success but in general, it fails to capture power" and it may even become negative (Napel and Widgren 2004: 524). A fundamental critique has come from Braham and Holler (2005) who deny the possibility of a preference-based power index on the ground that it is incompatible with "a fixed core of meaning of power", i.e., the basic notion of power as a generic ability.

This paper contains our responses and is organized as follows. Section 2 describes the logic of the SPI. In Sect. 3 we, first, present the arguments leading Braham and Holler to deny the possibility of a preference-based power index. We then demonstrate why these arguments are not convincing. Section 4 deals with the argument put forward by Napel and Widgren that the SPI is not a true power index since it confuses power and luck. Section 5 addresses the question of whether the SPI can become negative. Section 6 is concerned with the Felsenthal and Machover proposition that the SPI is nothing but a modified Banzhaf index. Section 7 concludes the paper and presents our outlook.

2 The Strategic Power Index

As discussed by Steunenberg et al. (1999), the strategic power index approaches power as a player's ability to affect the equilibrium outcome in a game. The basic intuition is that the stronger a player's influence on the outcome under a specific game form, the more powerful this player is. The index is based on several elements: the modeling of a decision-making process using the tools of noncooperative game theory, the definition of a state space (outcome space), a distribution of state variables, and the use of an index to measure power.

The first element of the approach is the development of a game-theoretical model of a decision-making process, which includes various structural elements such as the set of players, their action sets, the possible sequence of moves, the distribution of information, the set of outcomes, and an outcome function mapping the space of strategy profiles into a set of outcomes. These elements define what is called a game form.

The second element is to define the preferences of the players. Let $n \in N$ be the number of players in a game form π and $X \subseteq \Re^m$ an m-dimensional and finite outcome space. For this space players are assumed to have Euclidean preferences which can be characterized by player i's ideal point $x_i = (x^1{}_i, x^2{}_i, \ldots x^m{}_i)$.[6] Let $q \in X$ denote the status quo, that is, the hypothetical state of affairs before the start of the decision-making process. This can be the current policy, or the situation without such a policy. We call a combination of a particular ideal point for each player and the status quo a 'state of the world', which will be denoted as $\xi = (x_1, x_2, \ldots x_n, q)$.

The third step is to feed the 'state' variables into the game form, π, based on some distribution. Combining a specific state, ξ, with the game form, leads to a specific game with some (unique) equilibrium outcome $x^\pi(\xi)$.[7] In this context, each particular state of the world is assumed to be the instance of a random variable $\underline{\xi} = (\overline{x}_1, \overline{x}_2, \ldots \overline{x}_n, \overline{q})$. In order to assess how well a player perform in a game form, we determine the expected distance between the equilibrium outcome and the player's ideal point for all possible configurations of preferences and the status quo, or states of the world. The expected or mean distance between the equilibrium

[6]For a more general version of the SPI, it is only necessary that X is some metric space, i.e. a space on which a metric (distance function) is defined, which, for every two points in X, gives the distance between them as a nonnegative real number. Such a metric space must satisfy the axioms of symmetry, positive definiteness and triangle inequality. The most familiar metric space is the (one- or multidimensional) Euclidean space which we assume in this paper. The Euclidean space is translation and rotation invariant and stretching, shrinking or mirroring at the origin does not alter the SPI.

[7]At this point we focus on a unique equilibrium outcome only for expositional convenience. The strategic power index can also be applied to games for which multiple equilibria exist. If the game does not have a unique equilibrium, but multiple equilibria, the simple Euclidean distance can be replaced by the average Euclidean distance, i.e. the sum of the Euclidean distances between each equilibrium outcome and the player's ideal point for all equilibria in a particular state of the world, divided by the number of equilibria.

outcomes for some game form, π, and player i's ideal point is given by

$$\Delta_i^{\pi} = \int \delta_i^{\pi} f\left(\overline{\xi}\right) d\left(\overline{\xi}\right)$$

where

$$\delta_i^{\pi} = \sqrt{\sum_{k=1}^{m} \left(x^{\pi}\left(\overline{\xi}\right)^{k} - \overline{x}_i^{k}\right)^{2}}$$

is the Euclidean distance between the equilibrium outcome of the game and the ideal point of player i in any particular state of the world, and $f\left(\overline{\xi}\right)$ is the density function if $\overline{\xi}$ is a continuous random variable.

The mean distance, as expressed by Δ_i^{π}, provides information on how well a player is doing in the context of a game form. This distance allows us to assess a player's power vis-à-vis the other players: all other things being equal, a player is more powerful than another player if the expected distance between the equilibrium outcome and its ideal point is smaller than the expected distance for the other player. This measurement forms the basis of the proposed power index.

The last step of the approach concerns the development of an index. In order to distinguish 'power' from 'luck', which are both contained in the measurement of expected distances, we need some standardization. This is done by comparing the 'performance' of a player with that of a dummy. A dummy is defined as a player with preferences for the same space as actual players, but who does not have any decision-making rights in the game form. Moreover, dummy player's preferences are independent of the other players. As a consequence this player or his/her preferences do not matter for the outcome of the game. The dummy only experiences some equilibrium outcome that is set by the other players. Sometimes the dummy is 'lucky' in having an ideal point that is close to the equilibrium outcome. However, in other 'states of the world' the dummy may be less fortunate and encounter a policy outcome that is quite different from its preferred option. Consequently, the mean distance found for this player represents a minimum value that can be associated with a 'powerless' player and provides a baseline above which we can speak of power.

For a dummy player, d, the expected distance between his ideal point and the equilibrium outcome of a particular game based on game form π can be defined as Δ_d^{π}. The (absolute) strategic power index for a player i can now be defined as[8]:

[8]The relative power of player i can be defined as $\tilde{\Psi}_i^{\pi} = \frac{\Delta_d^{\pi} - \Delta_i^{\pi}}{\sum_{j=1}^{n}\left(\Delta_d^{\pi} - \Delta_i^{\pi}\right)}$. The relative power scores of all players add up to 1.

$$\Psi_i^\pi = \frac{\Delta_d^\pi - \Delta_i^\pi}{\Delta_d^\pi} = 1 - \frac{\Delta_i^\pi}{\Delta_d^\pi}.$$

This index lies in the interval [0,1] and increases with the power of player *i*. The expected distance for a player that is 'powerful' enough to dictate the outcome of a game under any preference configuration would be zero, leading to a corresponding value for the index of one. By contrast, if a player has an effect on the outcome of a game, similar to that of the dummy player (which, by definition, is 'powerless'), the expected distance for this player is the same as for the dummy player, leading to a corresponding index value of zero.[9]

Based on this index, there is a natural way to approach the status quo bias of a game form, that is, the extent to which players are unable to act and to pull a new policy away from the current state of affairs. For a specific game form, that status quo bias can be measured by the expected distance between the equilibrium outcome and the status quo, which is defined as Δ_q^π. Substituting this value for the expected distance found for a player in the strategic power index, we get

$$\Psi_q^\pi = \frac{\Delta_d^\pi - \Delta_q^\pi}{\Delta_d^\pi} = 1 - \frac{\Delta_q^\pi}{\Delta_d^\pi},$$

which is called the *inertia index*. A value of one for this index means that under some game form the status quo always prevails. The smaller the value for the index, the more players are able to move the equilibrium policy away from the status quo.

3 Impossibility of the SPI?

In this section, we deal with the critique put forward by Braham and Holler (2005) who argue that a preference-based power index is impossible. We, first, present the main argument leading Braham and Holler to deny the possibility of a SPI. Next, we discuss its main shortcomings.

3.1 A "Core Theorem of the Measurement of Power"

Braham and Holler want to bring the "semantics of power into the centre of the debate about how to measure power" (Braham and Holler 2005: 139). By referring to the philosophical semantic analysis of power they take the notion of a "generic

[9] Since the ideal points for each player are independent random variables, the equilibrium outcomes can never be systematically biased *against* the interest of a particular player, and, therefore, no player can fare worse than the dummy player. Thus, the proposed index can never become negative.

ability to effect outcomes" to be "the natural 'fixed core of meaning' of power" (Braham and Holler 2005: 145). In their words:

> *If* player *i* wanted a particular outcome or set of outcomes and that *i* has an action (or sequence of actions) such that the performance of these actions under stated or implied conditions will result in that outcome or set of outcomes and would not result if *i* would not perform this action (or sequence of actions), *then* player *i* would perform this action (or sequence of actions) and the specified outcome or set of outcomes would obtain. That is, *i* is essential or non-redundant for an outcome or set of outcomes (Braham and Holler 2005: 145).

In the absence of that player's intervention the state of the world would be different (Braham and Holler 2005: 145). Regarding simple games, one would speak of a swing (Braham and Holler 2005: 145, n. 8) or a player being decisive or pivotal.

From the definition of power as a capacity or potential to affect outcomes Braham and Holler conclude that a measure of power cannot accommodate any reference to the preferences of the players with respect to affecting outcomes. A power ascription is, first, categorical, second, "leaves the matter of what *i* wants undefined" and, third, "does not say how much power *i* has, only that there exist circumstances in which *i* is non-redundant for the outcome; a measure of power—power index—aggregates these ascriptions of non-redundancy in some way" (Braham and Holler 2005: 145–146).

The central claim of the article is formulated as 'Core Theorem of the Measurement of Power', which is, as the authors concede, not a theorem in the formal sense of the term, but rather "a kind of conceptual impossibility result that is germane to the theory of power generally" (Braham and Holler 2005: 138). The 'theorem' is stated in the following way:

> Core Theorem of the Measurement of Power: If power is the ability of *i* to affect an outcome, then a measure of *i*'s power must exclude any reference to *i*'s preference (behavioural content) with respect to affecting that outcome (Braham and Holler 2005: 146).

Three reasons are given for this statement:

(1) being disinclined to do something does not imply the inability to do it;
(2) psychological states such as desires and wants are not normally applied to the concept of ability; and
(3) the *exercise* of an ability is not to be conflated with its *possession* (Braham and Holler 2005: 146).

Braham and Holler are of the opinion that a preference-based power index such as the SPI violates these three conditions: it conflates disinclination with inability (Braham and Holler 2005: 146–148); redefines the game form, since a 'phobiafied' strategy, i.e. a strategy which is not rational being chosen, cannot be considered a strategy at all (Braham and Holler 2005: 148–150); commits the so called *exercise fallacy* by conflating the "*possession* of a disposition (*having* power) with its *exercise*" (Braham and Holler 2005: 151).

With regard to a game theoretical setting, Braham and Holler 'derive' a corollary of their theorem, which

> states that a player's power resides in, and only in, the strategies available to her given by the *game form* and not in the way that she plays the *game*. This implies that power is a value-independent concept. The upshot is that the Core Theorem renders unintelligible any attempt to formulate a measure of power in terms of the equilibrium of a non-cooperative game – the very idea of strategic power indices. Put bluntly, assessing how a player may play a game does not help us answer such questions as 'Is Smith more powerful than Jones?' or 'What is the extent of Smith's power?' because power concerns what players may be *able to do*, not the actions they *may* or *do* take (Braham and Holler 2005: 139).

Interestingly also to Braham and Holler power is linked to the game form and the strategies available to players. The SPI gathers information on the success of the various, available strategies, which are embedded in the game form, by comparing the outcomes for a distribution of states of the world. Following a broad and general distribution basically levels out the effect of specific values related to preferences and policies, which seem to be the main reason of Braham and Holler's objection.

3.2 Is There a Fixed Core of Meaning of Power? (Pitfalls of Essential Definitions)

The essay of Braham and Holler is an exercise in semantics. They concede that they are "making liberal use of the philosophical semantic analysis of power conducted" (Braham and Holler 2005: 139), but they add: "It must not, therefore, be thought that we are refreshing old philosophical debates. Rather, we are bringing the *semantics* (italics added) of power into the centre of the debate about how to measure power" (Braham and Holler 2005: 138). In fact, they claim having formulated the 'right' ('true') definition of power, with 'general ability to affect outcomes' constituting its essence or intrinsic fundamental nature. In the philosophy of science those definitions are called 'essential definitions'. The problems associated with essential definitions are well known (Popper 1960, Chapter I.10): Is there one, and only one, notion of power? How do we know that 'general ability' is the essential property of power? How can we evaluate the definition in terms of the truth or falsity of the description given by it? Referring to "what we customarily mean by ability" (Braham and Holler 2005: 144) is a doubtful criterion, raising more questions than solving ones.

We should try to avoid converting substantial problems in purely semantic (verbal) ones, since this paves the path for endless discourses. We should reject the view that we should aim at and can obtain ultimate explanations by looking for essences. Following the path of methodological nominalism, definitions such as 'power is a generic ability' should be read from right to left, as an answer to 'What *shall* we call a generic ability in a game form?', and not from left to right as an answer to 'What *is* power in a game form?' Accepting this rule, one would be rather reluctant in conducting an 'analysis of power per se', as done by Braham and Holler (2005: 154). Since we do not have a criterion for figuring out what 'power per se' actually is, it seems reasonable to take an instrumental stance to the definition and to ask, 'Which definition is helpful in answering scientific questions?', and: 'Why are we interested in a definition of power?' The answer clearly depends on where we want to use the term. Several possibilities come to mind: If power is part of a theory, the explanatory power of the theory might depend on the definition. For normative statements, the workability or the empirical relevance of a concept of power in the sense of 'What people are really interested in' might be decisive. From this perspective one might ask: Of what interest is it to know what a player is able to do, if it is not rational to do it? Why should we be interested in the potential or capacity of an action to alter outcomes if this does not is in accordance with equilibrium behavior?[10]

Our position is that power can and should be defined in several ways depending on the research question. In some contexts it might be useful to define power the way Braham and Holler did, i.e. applying the criterion of decisiveness, in others it is better to follow the SPI approach, which relies on the criterion of success. In Sect. 3.3.3, we will show that in take-it-or-leave-it committees—these are the voting bodies to which the SPI is applied—the criterion of success is the better measure of power. In the next sections we illustrate the relevance of the above arguments by analyzing some well-defined games.

3.3 Thinking Strategically vs. 'Analysis of Power Per-Se' or: Why the Inclusion of Preferences Is Necessary

Strategic interactions arise in two forms. The first is sequential. Here, players make alternating moves whereby earlier moves are observable to those choosing later. In a simultaneous game, players act at the same time in ignorance of the other player's current actions (game of imperfect information).

3.3.1 Constitutional Choice[11]

Consider three legislators, A, B, and C, who must vote in alphabetic order under a majority rule, on whether to increase their own salaries (see Ordeshook 1992: 41f.). Each legislator prefers to receive the pay raise, but each realizes that the constituents will not be pleased with a legislator voting to increase his own salary. There are four possible outcomes (see Ordeshook 1992: 41):

o_1: The raise passes, but the legislator votes against it.
o_2: The raise passes, and the legislator votes for it.
o_3: The raise fails, and the legislator votes against it.
o_4: The raise fails, but the legislator votes for it.

[10] See Barry's critique of the Shapley–Shubik index (Barry 1980).

[11] This part is from Schmidtchen and Steunenberg (2002: 208–210).

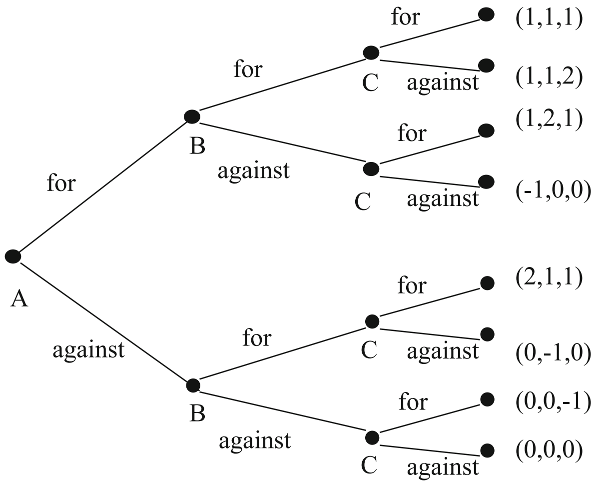


Fig. 1 The pay-raise game in extensive form

Let u denote utility, then the preferences of the legislators are summarized by the following numbers

$$u_{\mathrm{i}}(\mathrm{o}_1) = 2, u_{\mathrm{i}}(\mathrm{o}_2) = 1, u_{\mathrm{i}}(\mathrm{o}_3) = 0, u_{\mathrm{i}}(\mathrm{o}_4) = -1, \quad \text{with } \mathrm{i} = \mathrm{A}, \ \mathrm{B}, \ \mathrm{C}.$$

Figure 1 represents this voting situation in extensive form, where the terminal nodes are associated with the payoff 3-tuple (u_A, u_B, u_C).

A game form analysis of this voting game, i.e. neglecting the payoffs of the players, which applies 'general ability to affect outcomes' as indicator of 'power per-se' would reveal that each player has power: If two players were to vote differently the third one is decisive, he decides which state of the world obtains—pay-raise or status quo. Since this result holds for each player, traditional power indices would assign equal power values to each of the three players. However, the game is a sequential one, which matters a great deal.

Assume society consists of our three players who had to choose, say, unanimously the game form underlying the game that is to be played afterwards.[12] Which game forms are candidates for getting unanimous support?

From the point of view of power as a 'general ability to affect outcomes' the game form of Fig. 1 is a candidate. Power seems equally distributed—the sequential order of play does not matter. We doubt that players are so stupid not to see that the order of play is highly relevant. The legislators must vote alphabetically—player A is moving first, player B moving second and player C moves last. Player A

[12]This is a traditional constitutional choice problem (Buchanan 1990). On the constitutional level, society must choose the rules (choice of rules) that govern decision-making on the post-constitutional level. On the post-constitutional level, choices have to be taken within the rules decided upon on the constitutional level.

	A for		A against	
	B for	B against	B for	B against
C for	1, 1, 1	1, 2, 1	2, 1, 1	0, 0, -1
C against	1, 1, 2	-1, 0, 0	0, -1,0	0, 0, 0

Fig. 2 Pay-raise game in strategic form

enjoys a first mover advantage, which can be seen if we derive the subgame perfect equilibrium. The unique subgame perfect equilibrium is: (against, for, for). Player A receives the highest payoff (2), whereas the other players must content themselves with their second best outcomes. Note that if B and C were to change position the outcome of the game would not be affected. Clearly, each player would prefer to occupy the first mover position. A sequential game form such as in Fig. 1 would only have a chance to be chosen on the constitutional level, if uncertainty exists with respect to the first mover position. In such a case, each player must form beliefs about his position. With a perfect veil of ignorance these beliefs would be identical, leading to identical expected utilities for the players, given the majority rule $m = 2$. Similar calculations are required for the $m = 3$ and $m = 1$ rules.

Having done all these computations, the players can choose, on the constitutional level, which rule is best for them given the sequential order of play. But why should a sequential game be chosen at all? On the constitutional level players are free to choose a simultaneous game, which would change the structure of information of the pay-raise game dramatically.

Figure 2 portrays this game in strategic form (see Ordeshook 1992: 45); ordering of the payoffs (player C, player B, player A).

In contrast to the game portrayed in Fig. 1, where players A, B, C have 2, 4, 16 strategies, respectively, in the simultaneous game the strategy sets B and C are identical to A's—to vote for or against. Here, player C cannot condition on the choice of player A or B and B cannot condition on player A's choice. This game has four Nash equilibriums: (A for, B for, C against), (A for, C for, B against), (A against, B for, C for), (A against, B against, C against). Thus, in this section we reached a conclusion similar to that in the previous Section. A constitutional analysis, which restricts itself solely to the analysis of game forms, would be incomplete.

3.3.2 Inferior Players

Next, consider a 3-player simple game where the only winning coalitions are the grand coalition ABC and the two coalitions AB and AC (this example is from Napel and Widgren 2001: 213; Widgren and Napel 2001: 1–2). Looking at this game as a

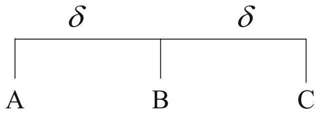


Fig. 3 Preference constellation

coalitional form game the Banzhaf and Shapley–Shubik power vectors are $(\frac{3}{5}, \frac{1}{5}, \frac{1}{5})$ and $(\frac{2}{3}, \frac{1}{6}, \frac{1}{6})$, respectively.

From the point of view of non-cooperative game theory—following Napel and Widgren (2001: 213)—the game can be looked at as a sequential game, in which A makes, after flipping a coin, an ultimatum offer to B, asking for approval in return for an only marginal (and in the limit non-extent) concession to B's interest. A rational player B would have to accept the proposal, if a blocking coalition BC cannot be formed. B knows that if he rejects the proposal, A would move to C, who serves as a perfect substitute in forming a winning coalition. A similar reasoning holds in the case in which A makes the ultimatum offer to C. Thus, we would conclude, contrary to what power measures based on coalitional form games indicate, B and C are powerless in this game. Napel and Widgren (2001: 213–14) call players that are robbed of their power commonly associated with their swing inferior players.

Indeed, application of the machinery of the strategic power index shows that B and C are powerless. Consider a policy space with three possible outcomes and identical distance, denoted δ, between two neighboring outcomes; player set {A, B, C} and D as dummy player. This player is not a true player but rather an outside observer. The ideal points are uniformly distributed on the policy space. Figure 3 shows one of the feasible preference constellations.

Translating the notion of an ultimatum game to our setting means that, whatever the distribution of ideal points (preference profile) of players A, B, C, the policy outcome always corresponds to A's ideal point. Thus, A's power score $\Psi_A = 1$. Since we assumed that the probability distribution of D's ideal points is the same as those of B and C, D's expected distance equals those of B and C: $\Delta_B = \Delta_C = \Delta_D$. Thus, $\Psi_B = \Psi_C = 0$.[13] The conclusion is that traditional power indices assign power scores to players without taking into account their position as inferior players. Thus, they neglect a factor that may be highly relevant from a player's point of view.

3.3.3 Agenda Setting Versus Veto Power[14]

Related to the first mover issue analyzed above is the problem of agenda setting power and veto power. Agenda setters are first movers, but not every first mover

[13] Note the difference to the strict power index approach favored by Widgren and Napel, where A is not treated as a 'pure' ultimatum player. Whereas A's SPI score is 1, the strict power index is 5/7. However, according to both indices B and C are powerless.

[14] This part is based on Schmidtchen and Steunenberg (2002: 212–214).

is an agenda setter. Classical power indices are not able to analyze and evaluate these distinctive types of power since all players are simply veto players. There are only two subsets of players, a 'winning' and a 'losing' one. Braham and Holler acknowledge that those indices are insensitive to the strategic aspects of power relations (Braham and Holler 2005: 141). They even illustrate this feature by two elementary examples. In one of their examples they consider a committee of seven players with each player having one vote and a 5/7-majority rule (Braham and Holler 2005: 142–143). They further assume a preference configuration ranking the players in a one-dimensional policy space and a proposal falling from heaven. They show that not all coalitions are rationally feasible, and that not every swing will be exercised by a rational agent (Braham and Holler 2005: 142). However, despite acknowledging the intuitive appeal of the critique encapsulated in the examples, they are still having the opinion that "it is fundamentally mistaken. The reason hinges on a conceptual issue: what we mean by a power ascription" (Braham and Holler 2005: 143). A substantial problem is converted into a verbal one! In most of the committees there are agenda setters. Furthermore, it is well known that a specific type of power, different from the power of a veto player, is associated with the position of an agenda setter as a first mover. A power concept that systematically neglects the sequential structure of collective decision-making is unable to measure this type of power and to address the problems associated with it.

To show why traditional voting power indices do not represent the distribution of power between an agenda setter and several veto players in a satisfactory and meaningful way, we choose, as a simple example, a decision-making procedure used in the European Union. With regard to legislative decision-making, the EC Treaty initially provided only for the unanimity version of the consultation procedure. This procedure allowed the Commission to propose new regulations or directives, which are subjected to unanimous consent by the Council. The latter implies that, in fact, each Council member has the right to veto the Commission's proposal. The European Parliament only needs to be consulted in this procedure. Since the Council can adopt a proposal regardless of the position Parliament takes, Parliament does not play a significant role and thus will not be discussed further.

Now assume that policies can be represented by a one-dimensional (left-right) outcome space and players have Euclidean preferences. In addition, assume that players have perfect and complete information. The Commission selects a proposal, which is then decided upon by the Council members. For our argument on the usefulness of voting power indices, we assume that Council members are not allowed to add new proposals to the agenda or to amend the Commission proposal. The interactions between the Commission and Council members now resemble the well-known agenda-setter model of Romer and Rosenthal (1978, 1979).

Figure 4 presents a preference configuration that may occur for the Commission, which is conceived as a unitary actor, and a five-member Council. In this figure V_i and C denote the most preferred or ideal points of Council member *i* and the Commission, respectively, and $V_i(q)$ stands for member *i*'s point of indifference to the status quo q. The Commission, C, has a more progressive preference than most Council members, V_i. Nevertheless, the leftmost Council member, V_1, holds

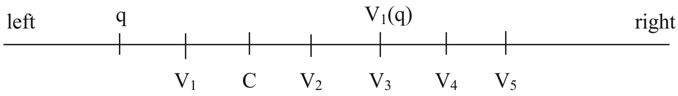


Fig. 4 Preferences of the Commission and the Council Members

an even more extreme position. Given a status quo to the left of these players, the Commission will propose a measure that is equivalent to its own most preferred point. Since all Council members prefer this point to the status quo, the proposal will not be vetoed. So, in equilibrium, the outcome of this game is a legislative policy $x = C$.

In this context, all players have to approve a measure, and no measure can be taken without the support of each one of them. Each (last) player has the same probability of being pivotal, and each player is necessary to form the (minimum) winning coalition of all players. The Shapley–Shubik, Banzhaf, Johnston and Holler indices therefore allocate power values of 1/6 to each player. These individual scores would suggest that the Commission is as 'powerful' as the Council members. The aggregated score of the Council would even be 5/6, which implies that the Council would be *more* powerful than the Commission.

Adding together the scores of individual Council members to calculate the power of the Council leads to what we call an *aggregation bias*. This bias is the result of the fact that, in interbody analyses of voting power, the members of separate decision-making bodies are treated as if they were the members of a single committee. However, in the game as discussed, a proposal must be approved by both the Commission and the Council, regardless of the voting rule the Council uses to reach a collective decision. If the Commission does not belong to a coalition, then this coalition is not a winning coalition. Both players can be regarded as necessary players. Therefore, one would expect that both actors have power values of 1/2, and not 1/6 for the Commission and 5/6 for the Council. The bias, as revealed by these numbers, leads to an exaggeration of the Council's abilities and an understatement of the power of the Commission. The reason is that the abilities of these players to affect the equilibrium outcome differ: the Commission can take the initiative and draft a proposal, while Council members can only approve or reject this proposal. Council members may restrict the Commission's policy choice, but they cannot set the final proposal. The Commission enjoys discretion in choosing a new policy, which makes it more 'powerful' than the traditional indices indicate.

In addition, the power value of the Council, in a game with the Commission, is independent of the number of Council members. The individual values are only relevant to assess each member's power in shaping a Council decision and not a decision that has to be taken by several 'institutional' actors, including composite decision-making bodies.

3.4 Strategic Power: Ability of Being Successful

The SPI measures a player's ability/capacity/potential (whether generic or not) on average to influence (affect) as a member of a voting body the equilibrium outcome of a voting game or, in other words, the ability/capacity/potential 'to get what you want' by incentivizing as a member among other members of a voting body an agenda-setter to present proposals which approach as close as possible the preferences of the respective player. It is an indicator of average success of affecting *equilibrium* outcomes.

This potential to affect equilibrium outcomes is determined by the game form, the state space and state variables (which are random variables). Taking into account the preferences of the players serves the purpose of determining rational behavior and to derive the equilibrium in a specific game.[15] Since the sole sources of power are the game form, the state space and the state variables we can fully subscribe to Braham and Holler's statement:

> Ordinarily speaking, a 'power' ascription refers to a person's ability: what a person is able to do. In the game theoretic context that we are discussing, the ability in question is to effect outcomes (i.e. 'force' or 'determine' outcomes) of the game. That is, a player has a strategy that, if chosen, will make a decisive difference to the outcome. *This basic definition is the same for a power index based upon a simple game and one that is ostensibly based upon a non-cooperative game* (italics added). The difference lies in the specification of the ability. In a simple game, the ability is turning a winning coalition into a losing coalition or vice versa, thereby being decisive for the acceptance or rejection of a bill, while, in a non-cooperative game, the ability is specified in terms of shifting the equilibrium in one's own favour (Braham and Holler 2005: 143).

Note that in both models of a decision-making procedure the veto-players have identical action sets: they can either reject or accept a proposal. But only in the non-cooperative game setting players are assumed to act rational, i.e. choosing that action which leads to the better individual payoff.

It depends on the decision-making rule whether or not a player is *decisive* as for the equilibrium outcome. With a unanimity rule each veto player is decisive in the sense that the rejection or acceptance of a proposal always, i.e. whatever the preference configuration, depends on the action chosen. With a rule of simple majority there are sometimes preference configurations in which the equilibrium outcome of the game, i.e. either the status quo or, if there is a proposal, its content, crucially depends on the action of a player; but sometimes the equilibrium outcome is determined irrespective of the action chosen by a player. Nevertheless, in the latter case still distances between the ideal points and the equilibrium outcomes can be calculated and they are included in our power measure.

From the discussion above it should be obvious that, contrary to what Braham and Holler (2005: 147–148) believe, taking into account the state space and state variables in measuring a player's power does not mean conflating disinclination

[15] Of course, our approach can also be applied to games of incomplete information, which would require making assumption regarding the possible types of players.

with inability. What the SPI measures is simply the ability/capacity/potential of a *rational player* to affect an (equilibrium) outcome, which is a subset of all possible outcomes. Finally, contrary to what Braham and Holler (2005: 150–152) believe, we do not conflate the possession of power with its exercise thereby committing the so-called exercise fallacy. What the players, the agenda-setter and the veto-players, do is exercising rational behavior. Whether or not, for example, a veto-player affects the equilibrium outcome depends on the decision-making rule and the rational behavior of all other players.

Meanwhile even adherents of the traditional power index approach question that there is only one notion of voting power, namely decisiveness. They realize that the notion of 'satisfaction' or 'success', "that is, focusing on the likelihood of having the result one voted for irrespective of whether one's vote was crucial for it or not" (Laruelle et al. 2006: 186) is a meaningful notion of "voting power" and might be more relevant than decisiveness from the voters' point of view (Laruelle et al. 2006: 189; Laruelle and Valenciano 2008). Whereas in so-called bargaining committees decisiveness is the adequate notion of power, in so called take-it-or-leave-it committees—these are the committees in which the set of players "is entitled only to vote for or against proposals submitted to it by an external agency" (Laruelle and Valenciano 2008: 53)—success is the better one (Laruelle and Valenciano 2008). We agree but there remains still a difference to our measure of success: Laruelle et al. measure success by a probability, whereas we take the expected distance between a player's ideal points and the equilibrium outcomes.[16] But irrespective of this difference, what Laruelle et al. (2006: 201/203) conclude is worth to be quoted:

> Perhaps the fascination raised by the notion of 'power' has caused a distortion of focus in the field. It can be argued that decisiveness seems intuitively closer to the notion of 'power' than that of success, but this does not grant greater credit to recommendations based on this interpretation. In other words, the relevant question is not what notion is closer to the intuitive idea of 'power', but is a more adequate basis for normative recommendations. And as a base for normative recommendations (e.g., in connection with important issues, as that of the most adequate voting rule in a 'take-it-or-leave-it' committee of representatives) it seems more relevant the notion of success than that of decisiveness.

The upshot of these deliberations is that a decision-making process can be modeled in several ways: as a simple game, using a coalitional or purely probabilistic approach, or as a non-cooperative game. Whether the one or the other is superior depends upon the question to be addressed *and* whether the nature of the decision-making process—for example, the sequential moves of the players, the inter-body decision-making, or the possibility to vote strategically (Schmidtchen and Steunenberg 2002: 206–214)—is adequately captured. To paraphrase Braham and Holler: "Here lies the heart of the problem" (Braham and Holler 2005: 144).

[16] Another difference is worth to be mentioned: Whereas in the Laruelle et al. model proposals are submitted by an external agency, the agenda setter in our model is a player, thinking strategically.

4 The SPI: Confounding Power with Luck?

In a much-cited paper entitled "Is it Better to be Powerful or Lucky?" Brian Barry presented the following formula: success = luck + decisiveness (Barry 1980: 338). Although Barry had not been concerned with non-cooperative voting games and, moreover, defined the terms as probabilities, the logic of this formula applies in a modified way to the SPI as well. The modification consists in substituting, first, probabilities by distances between ideal points and equilibrium outcomes, and, second, decisiveness by strategic power. As for the latter substitution, recall that power has nothing to do with decisiveness but refers to the ability of getting desired (equilibrium) outcomes. In order to level out the effect of luck, we focus on the average or expected ability.

This procedure has been criticized by authors who are in favor of power indices quite similar to ours. For example, Napel and Widgrén propose—as we do—a unified framework for measuring power as determined by spatial preferences, strategic agenda setting and decision-making procedures (see Napel and Widgren 2004). Thus, they do not deny the possibility of a preference-based power index. However, they claim that the framework underpinning the SPI leads to a *strategic success index*, rather than a strategic power index. In their view, SPI measures "the ability of a player to make a difference in the outcome", i.e. power, only under very special circumstances (Napel and Widgren 2004: 524): "Only for particular distribution assumptions is luck 'leveled out' by taking averages" (Napel and Widgren 2004: 524), and: "Unless one regards average success as the defining characteristic of power (which neither Steunenberg et al. nor many others do), taking expectations will only by coincidence achieve what Steunenberg et al. aim at, namely' to level out the effect of 'luck' or a particular preference configuration on the outcome of a game'" (Napel and Widgren 2004: 524). Napel and Widgren concede that the SPI "is a good measure of average success but, in general, it fails to capture power" since the SPI confounds luck with power (Napel and Widgren 2004: 524).

We will discuss this critique in turn. Consider Fig. 4, which can be used to illustrate the importance of distinguishing 'power' from 'luck'. The equilibrium outcome of the game is $x = C$, that is, the most preferred position of the Commission. This outcome seems to be more favorable to Council member 2 than member 5, since the distance to V_2 is less than the distance to V_5. Is member 2 therefore also more powerful? Both players have the same abilities to affect the outcome, that is, to veto the Commission proposal. So, from this perspective, there is no difference in power. Nevertheless, the outcome is closer to member 2's preferences. This indicates that member 2 is more 'lucky' than member 5. Having a preference that lies close to the equilibrium outcome of a particular game does not necessarily mean that this player is also 'powerful'. Similarly, one may question whether Council member 1 is more 'powerful' than the other Council members, since this player defines the boundary, $V_1(q)$, where the Commission can no longer select its ideal point, should this player move to the right. If any other player can also occupy the position of this member, or the status quo can be located at any other point along the policy dimension, Council member 1 is just more 'lucky' than the others. Following Barry (1980), we regard (in this specific game) a player's *success*, which is defined as the extent to which the outcome of the decision-making process corresponds to its ideal point, as the composite effect of 'power' and 'luck'. Part of a player's success is therefore based on 'luck', the other part is due to the 'power' a player exerts.[17]

Whereas power can be associated with a player's ability to affect the final outcome [which is basically a matter of the rules of the game telling us *who* can do *what* and *when* and who gets *how much* when the game is over (see Binmore 1992: 25)], 'luck' is related to the preferences of the players and the location of the status quo, which are assumed to be exogenously determined. The latter can be illustrated by the role of the Commission in our example of the consultation procedure. The fact that the outcome of the game coincides with the Commission's most preferred point does not imply that the other players in the game are 'powerless'. This result depends on the preferences of the Council members and the location of q. A shift of V_1 to the left may, for instance, force the Commission to propose a policy $x = V_1(q)$. Thus, given the preference configuration, the Commission is 'lucky' that Council members have preferences that allow for the equilibrium outcome $x = C$. This clearly indicates that the *success* of a player in a given game is the combined result of *abilities* (defined by the rules of a game) and the specific *preference configuration*. To assess a player's power, a measure should be based on the former and not the latter.

To distinguish 'power' from 'luck', we propose a measure that is independent of the preferences of players in a specific game, which, together with the game form, determines the outcome of the game. This can be achieved by measuring a player's power under some game form with reference to the *mean* or *expected* distance between the equilibrium outcome and this player's ideal point for all possible combinations of players' preferences and all possible combinations of the status quo. In doing so, the power-luck confusion vanishes. The fact that our power

[17]Note the difference between our definition and Barry's definition, which has recently been given more precision by Laruelle and Valenciano (2008: 54–55, 58). In their view a player is successful *ex post*, i.e. once the players have voted on a given proposal, if he/she obtains an outcome—acceptance or rejection of a proposal—that he/she has been voted for. A voter has been decisive if he/she is successful and his/her vote was crucial (critical) to that outcome. Luck is simply success without decisiveness, i.e. a player's vote is irrelevant for the outcome. Thus, Laruelle and Valenciano interpret decisiveness, success and luck as binary variables.

Our definition of terms is more general than Laruelle and Valenciano's, first, in that it refers not only to veto-players but also includes the agenda-setter. Second, in our framework, a player is successful if his/her vote influences the content of the proposal such that the equilibrium outcome moves towards his/her ideal point (including the case in which the status quo remains). Contrary to Laruelle and Valenciano, in our framework a player can be more or less successful, since the distance between the equilibrium outcome and a player's ideal point can vary.

scores turned out to be sensitive to a change of the decision-making procedures (all other things being equal) gives further support to this conclusion.[18]

5 The SPI: Can It Become Negative?

Napel and Widgren claim that the SPI may become negative (Napel and Widgren 2004: 524; Napel and Widgren 2002: 9–11). We discuss two examples Napel and Widgren developed in support of their claim (Napel and Widgren 2002: 9–10; Napel and Widgren 2004: 524).

Consider a simple majority voting game with three players having equal voting weight and outcome space $X = \{-1, 0, 1\}$. Player 1's random ideal point, λ_1 is degenerate and always equal to 0, whereas the ideal points of players $i = \{2, 3\}$, λ_i, are uniformly distributed on X. The status quo is fixed on position 0. In only two out of nine states of the world $\xi = (q, \lambda_1, \lambda_2, \lambda_3)$ with $q = 0, \lambda_1 = 0$, and either $\lambda_2 = \lambda_3 = -1$ or $\lambda_2 = \lambda_3 = 1$ the status quo does not prevail. Since average distance is 2/9 for player 1 and 4/9 for both other players, player 1 appears to be the most successful and most powerful player. Napel and Widgren (2002: 10) conclude: "However, exactly the same equilibrium outcomes prevail when player 1's voting weight is reduced to zero, i.e. if he becomes a dummy player" (assuming that for an even number of players the status quo wins unless defeated by a majority). And they add: "According to Steunenberg et al.'s Strict (!) Power Index, he is still the most powerful player".

What is the reason for this seemingly strange conclusion? Given that the status quo is always q = 0 and the ideal point of player 1, λ_1, is supposed to be always at the status quo, the set up of the game implies a status quo bias. Therefore it is not surprising to see player 1 coming out as the most "powerful" player. This status quo bias still exists if player 1 has a voting weight of zero, since a majority, in fact unanimity, is needed to defeat it. Player 1 is clearly in both scenarios the most successful player, but not the most powerful. His/her superior performance is due to a restriction of the set of possible states of the world from 81, under the assumptions we would use to calculate the index, to 9, leading to luck for player 1 and bad luck on the side of players 2 and 3.

The SPI is normalized by the introduction of a dummy player. Contrary to Napel and Widgren's approach, this player is not a true player but rather an outside observer. In fact, by assuming player 1 being a dummy player Napel and Widgren transform the three-player game into a two-player game.

[18]Note again the difference between our approach and that proposed by Laruelle and Valenciano (2008: 58). They define the *ex ante version* of the three terms success, decisiveness and luck (irrelevance) using probabilities. The probability of a player being decisive is simply the difference between his/her probability of being successful minus the probability of being lucky.

We define a dummy player as “a player whose preferences vary over the same range as the preferences of the actual players, but that has no decision-making rights in the game” (Steunenberg et al. 1999: 348). Napel and Widgren claim that this definition of a dummy player is not always meaningful:

“What does it mean to ‘vary over the same range’ if the so-called actual players’ ideal points (to stay in a spatial voting framework) have different supports; e.g. $\tilde{\lambda}_i$ is uniformly distributed on [0, 1] and $\tilde{\lambda}_j$ has triangular distribution on [1/2, 4]?” (Napel and Widgren 2002: 11).

The answer is that the expression ‘same range’ refers to the range in the policy space in which the ideal points of *all* players can be distributed. In the example given by Napel and Widgren it is the range [0, 4].

Next, consider the second example developed by Napel and Widgren with the purpose to illustrate that, contrary to our view (Steunenberg et al. 1999: 349, n. 7), “equilibrium outcomes can be systematically biased against the interest of a particular player” (Napel and Widgren 2002: 11). If so, the SPI can become negative. In a group of four boys, the oldest one is the agenda setter and makes proposals as for what to do in the afternoon. Proposals have to be accepted by a majority of the remaining three boys. All boys have independent preferences, which follow the same distribution. According to the SPI framework the oldest boy as agenda setter is the most powerful player, and the SPI value is the same for the remaining players. There is a little brother of the oldest boy who is allowed to participate in the afternoon activities of the group but does not have a say in selecting the program. Regarding its preferences Napel and Widgren (2002: 10) make the following crucial assumption:

> It is plausible to assume that he does not always agree with his elder brother’s most desired outcome, but does so more often than with the others’ ideal alternatives. Mathematically speaking, let the ideal points of the two brothers be positively correlated. Then, the mean distance between the group’s equilibrium activity and its youngest member’s most desired recreation will be smaller than that of those group members who actually have their vote on the outcome.

These examples seem to suggest that the SPI is a rather strange construct: negative power—what sense does that make? But we should be rather careful in drawing such a conclusion: First of all, Napel and Widgren concede that “(f)or a measure of average normalized success this (negative power, the authors) makes sense” (Napel and Widgren 2002: 11): “it simply indicates that (a) player … is less successful on average than a neutral member of the decision body would be” (Napel and Widgren 2002: 11).[19] Second, note that a negative SPI is due and only due to the introduction of a dummy and not due to the internal logic of the SPI. The introduction of a dummy simply serves the purpose of transferring the players’ expected distances into a range of 0 and 1 (normalization). Of course, the SPI makes perfectly sense without such a normalization. Third, even if the power of a player

[19]Napel and Widgren present an example in which a player n always has a position “opposite” of his $n-1$ colleagues (Napel and Widgren 2002: 11).

after normalization turns out to be negative, the relative power of the players having voting rights is still correctly indicated.

However, we could also try an entirely different route of argument. Taking the idea of a veil of ignorance seriously—and this idea is at the heart of measuring a-priori power—one might well ask whether the examples presented by Napel and Widgren are actually to the point. Following Felsenthal and Machover (2001: 94), "to obtain an a priori strategic measure we must go behind a veil of ignorance: we must minimize the information built into the state space and the distribution of the state variables". The crucial question then is: what is the proper assumption regarding the distribution of the ideal points and the status quo? We feel that the veil of ignorance means that there is no information about the players' preferences and the status quo. Therefore, the principle of insufficient reason (or principle of indifference) requires assuming that the state variables are mutually independent and uniformly distributed on the state space.[20] From this point of view, a 'true' veil of ignorance (principle of insufficient reason or indifference) would imply two things, (1) independent distributed ideal points for all players having voting rights and the status quo and (2) the distribution of the dummy being identical to that of all other players. All the examples presented by Napel and Widgren violate this condition. Consider the little brother: despite the fact that little brother does not have a say in the game he is *not* a dummy player in the strict sense. He is simply lucky to have a preference closely related to that of the player that is most powerful. This is also the reason why the SPI would *not* assign power to Luxembourg, to take another example referred to by Napel and Widgren (2002: 11). Luxembourg is lucky having sometimes similar views with the other Benelux countries.

6 The SPI: A Banzhaf in Disguise?

In their comment on the 'Symposium Power Indices and the European Union' in the *Journal of Theoretical Politics* Felsenthal and Machover argue that strategic power is simply the Banzhaf power multiplied by a constant that depends on the shape of the state space (see Felsenthal and Machover 2001). If Felsenthal and Machover's conclusion were correct we would take this as support for our view that the SPI is a possible and reasonable measure of power. In the following we, first, present a sketch of the Theorem proven by Felsenthal and Machover, which is then followed by an evaluation of the results.

[20] We agree with Felsenthal and Machover (2001: 94) that this is not sufficient, "because the geometric structure of the state space itself also carries some information. In particular, any asymmetry of this space implies a bias in favour of some states and against others".

6.1 The Theorem

Consider a simple voting game ***W***. Let **S** denote a state space which is perfectly symmetric.[21] $X_i, \ldots X_n, Y, Z$ are independent random variables, all of which take their values in the state space. X_i, Y, Z stand, respectively, for the ideal point of player *i*, the state if a proposed bill will be passed, and the status quo (i.e. the state that continues to prevail if the policy proposal is defeated).

Let D_i denote the distance$|X_i - U|$ between *i*'s ideal point and the preferred state $U = Y$ or $U = Z$. This distance is a function of the random variables $X_i, \ldots X_n, Y, Z$ and can be regarded as a value of a random variable $D_i = f_i(X_i, \ldots X_n, Y, Z)$ which is completely determined by the simple voting game ***W*** and the joint distribution of the state variables (Felsenthal and Machover 2001: 93). As Felsenthal and Machover point out, "from the symmetry of **S** and the assumption that the X_i are independent and uniformly distributed on **S** it follows that the preferred state of each voter is equally likely to be nearer to **Y** than to **Z** as the other way around. Therefore each voter will vote 'yes' or 'no' with probability ½ and they will do so independently of each other—just as in the Bernoulli model underlying the *Bz* measure" (Felsenthal and Machover 2001: 94).

Now, let R and r, respectively, denote the greater and smaller of the two distances $|X_i - Y|$ and $|X_i - Z|$. Then the distance D_i can be defined as

$$D_i = (1 - p) \cdot R + p \cdot r,$$

with *p* the probability that *i*'s voting decision agrees with the outcome of the vote.

Using Penrose's theorem, which state

$$p = \frac{1 + \beta' (W)}{2},$$

with $\beta'[W]$ the Banzhaf, one can define the mean value of D_i, denoted $\Delta_i[W]$,

$$\Delta_i [W] = \frac{1 - \beta' [W]}{2} \cdot R + \frac{1 - \beta' [W]}{2} \cdot r,$$

for player *i*, and

$$\Delta_d [W] = \frac{R + r}{r}$$

for the dummy player. This gives

$$\Psi i [W] = \frac{R - r}{R + r} \cdot \beta_i' [W].$$

[21] Examples are: in the discrete case the set of vertices of a regular polygon or regular polyhedron; in the continuous case a circle or the surface of a sphere of some higher dimension.

Felsenthal and Machover (2001: 95) conclude:

> Thus $\Psi_i[W]$ is simply the *Bz* power of *i* multiplied by a constant that depends on the shape of **S**. Note, in particular, that in the simplest possible case, where **S** consists of just two points, *r* is clearly 0, so in this case $\Psi_i[W] = \beta_i'[W]$ exactly. In our view, this result vindicates the *Bz* measure: not for the first time, a new approach to the measurement of a priori I-power has, yet again, led to β'. It also suggests that the strategic measure proposed by SS&K is a natural generalization of a priori I-power, which allows the incorporation of additional information, and thus the study of a posteriori voting power.

Felsenthal and Machover believe that our method of measuring power is a promising candidate for a unified approach (Felsenthal and Machover 2001: 96). Since the SPI not only depends on the set of voters and the decision-making rule but also on the choice of the state space and the joint distribution of the state variables, there exists

> an enormous latitude for building into the model all kinds of information concerning the actual state of the world, the kinds of bill to be put to the vote and affinities or disaffinities between voters (Felsenthal and Machover 2001: 93).

6.2 Evaluation[22]

We welcome the Felsenthal and Machover approach. It forms a very interesting foundation of the approach presented here, which would allow the SPI to be fully characterized by the set of the axioms the Banzhaf is founded on. This axiomatic characterization would facilitate comparisons with other power measures. Although the theorem proven by Felsenthal and Machover provides for important insights into the logic of the SPI, several comments seem in order.

First of all, we agree that Felsenthal and Machover succeeded in reformulating the algorithm of the SPI as far as simple voting games are concerned. Simple voting games take the proposals to be voted upon as exogenously given. Thus, they can be treated—as in Felsenthal and Machover—as a random variable. However, the most important feature of the SPI namely the strategic interaction and the procedural constraints are not taken into account (see also Napel and Widgren 2002: 12–13; Napel and Widgren 2004: 524): the bills proposed in the SPI framework are not randomly chosen but are the result of strategic thinking along the subgame perfect equilibrium path.

To illustrate, consider Fig. 4. We know already, in equilibrium, the outcome of this game is a legislative policy x = C. Here, the outcome of a specific sequential game is partly due to the value of the random variables and partly the result of strategic thinking on the side of all players. It is natural to think about how to introduce this factor in the Felsenthal and Machover set up. One could take account of strategic thinking by restricting the domain of proposed bills in the

[22] We are particularly indebted to Stefan Klößner for several illuminating suggestions.

state space. The question is whether we can find some reasonable equivalent to the equilibrium concept used in non-cooperative game theory. On the other hand, one might conjecture that since proposals depend on the state of the world, including the ideal points of the Commission, and since the state of the world is a probabilistic variable, also the proposals are. In fact, one might even be tempted to apply the terms winning and losing coalitions in the context of a non-cooperative model of a decision-making procedure.[23] If a majority of the players vote in favor of a proposal then one could say that they form a 'winning coalition'. However, one should speak of a 'quasi-coalition' since, as Felsenthal and Machover (2001: 84) rightly mention, "the very term 'coalition', as referring to an arbitrary set of voters, is perhaps somewhat misleading, as it seems to imply conscious coordination". Moreover, contrary to traditional power indices, the SPI takes account of the fact that the propensity to present a proposal and its content depends on the composition of potential winning coalitions. In other words: The agenda setter is looking for a winning coalition such that the distance between its ideal point and the proposal (generating a winning coalition) is smaller than the distance between its ideal point and the status quo. If there is no such a winning coalition the agenda setter remains silent.

Second, the assumption of perfectly symmetric state spaces is very restrictive and reduces the applicability of the theorem considerably.[24] Note, for example, that the only symmetric one-dimensional state space comprises two points. Moreover, those sets are non-convex, which rules out interpreting a convex combination $\lambda x + (1 - \lambda)y$, with $\lambda \in [0, 1]$, as a compromise between x and y.

Third, perfect symmetry of a state space implies that yes/no decisions, interpreted as random variables, are stochastically independent. From this it follows that the distribution of ideal points, in nearly all cases, creates a 50 % a priori probability of a yes vote. Although the yes/no decisions of the players are stochastically independent, they are correlated in the following sense: They produce a distribution of the equilibrium outcomes for which the variance is considerably smaller than the variance of the status quo. The reason is that the proposals are less extreme than the status quo, and when player i accepts the proposal it is more likely that player j also does.

Fourth, probability p and, for that, the definition of distance D_i, implies that the outcome of the division agrees with the way player i voted (that is, the bill is passed (defeated) and i votes 'yes' ('no') (see Felsenthal and Machover 2001: 94).

Fifth, Felsenthal and Machover are of the opinion that the strategic power measure "is a natural generalization of a priori I-power, which allows the incorporation of additional information, and thus the study of a posteriori voting power" (Felsenthal and Machover 2001: 95). The notion of I-power is that of "power as

[23] As done by Widgren and Napel (2001) and Napel and Widgren (2004). See also the discussion in Sect. 3.3.2.

[24] Simulations indicate that in the case of perfectly symmetric state spaces the values of the SPI match those of the Banzhaf index, but they differ considerably for asymmetric state spaces.

influence: a voter's ability to affect the outcome of a division of a voting body—whether the bill in question will be passed or defeated" (Felsenthal and Machover 2001: 84). They argue that the notion of I-power "has essentially nothing to do with cooperative game theory or, for that matter, with game theory generally, as it is normally understood. According to this notion, voting behaviour is motivated by *policy seeking*. The action of a given voter does not depend on what other voters may be expected to do, let alone on bargaining and concluding binding agreements with them.

Each voter simply votes for or against a given bill on what s/he considers to be the merit of this bill; and the way s/he votes is independent of the decision rule. The passage or failure of a bill is here best regarded as a public good (or public bad), which affects all voters, irrespective of how they have voted on that bill" (Felsenthal and Machover 2001: 84).

We agree with Felsenthal and Machover that the SPI has nothing to do with cooperative game theory, but we disagree with Felsenthal and Machover's characterization of the SPI as not being in essence a game theoretic concept. First of all, although in a simple voting game the action of a given voter does not depend on what other voters may be expected to do, it depends on what the agenda setter has done. Second, the action of the agenda setter clearly depends on what s/he expects the other players will do (backwards induction). Third, application of the SPI approach is not restricted to simple voting games but has been applied to interbody decision-making (Steunenberg et al. 1999; Schmidtchen and Steunenberg 2002). Furthermore, in models allowing for the possibility of negotiating, amending or modifying proposals, forming coalitions and linking decisions on different proposals there is even more room for strategic considerations. Finally, the reformulation of the strategic power index, as presented by Felsenthal and Machover, is based on payoffs, since one cannot calculate differences without knowing the ideal points for all players. In fact, the constant with which the Banzhaf index has to be multiplied is a payoff measure. It is implicitly assumed that voters care about distances and that decisions are (rationally) determined by the distance of the ideal point from both the proposed bill and the status quo. These distances are utility measures.

In a comment on our 2002 article Moshe Machover takes up the issue that the distances in the state space can be interpreted as some kind of payoff contradicting the proposition that the SPI is not in essence a game theoretic concept (Machover 2002: 226–227). He thinks "that the contradiction is only apparent, not real" (Machover 2002: 227). This belief follows from his characterization of the model underpinning the calculation of the SPI as consisting of two distinct parts: "The first part is a decision rule, a so-called 'simple' game or 'simple voting game'. The second part consists of a state space and state variables (which are random variables). The decision rule operates in the conventional way: it tells us how the outcome of a division is determined by the way each of the voters vote. The second part of the model serves to model the motivation that leads each of the voters to vote in a particular way" (Machover 2002: 225). Although Machover is right in identifying the two distinct parts, he neglects the crucial fact that the second part not only serves to model the incentives of the voters but also the strategic choice of the agenda setter. Moreover, although Machover explicitly concedes that the geometry of the state space and the distribution of the state variables *is* game theoretic, belongs to non-cooperative game theory and the voters' decision "*may well* be based on a calculation of expected payoff" (Machover 2002: 227), he nevertheless sticks to his position that the model is not game theoretic: "The point is that in the case of I-power ... these motivations and payoffs are exogenous to the decision rule. This is precisely the situation in S&S's model: the decision rule resides in one part of the model, while the motivations and payoffs reside in the other part" (Machover 2002: 227). True, but we cannot see why this ubiquitous feature of models, i.e., consisting of several parts which are conceptually different, deprives the SPI of its game theoretic nature. A good model integrates different parts such that new insights are generated.

Finally, we share Moshe Machover's position

> that a correct method of measuring actual voting power should be organically connected with the method of measuring a priori power. The reason for this is that actual power is the result of a superposition of real-life factors (such as preferences) on the 'bare' decision rule itself. S&S's two-part model does precisely that; and when the contribution of the second part is reduced to nothing, the result is the Penrose measure (Machover 2002: 225–226).

However, reducing the second part to nothing would mean eliminating any strategic element in a power measure. The constant in the formula derived by Felsenthal and Machover simply disappears.

7 Conclusion

In this paper we have discussed the critique raised against the SPI as a preference-based power index. Overall, we find that the critique is unfounded: First, the proposition that the SPI is impossible results from playing with semantics. Second, the SPI does not confound power with luck, since taking expectations eliminates luck. Third, the SPI can become negative due and only due to the procedure of normalization. The SPI makes perfectly sense without a normalization. Moreover, taking the veil of ignorance concept seriously, it cannot become negative. Fourth, the attempt to show that the SPI is nothing but a modified Banzhaf and, for this reason, is not game theoretic should be welcomed, since it takes preferences into account and supports our claim for a unified approach to the study of a priori and a posteriori (actual) voting power. However, it neglects any strategic interaction and important procedural features such as, for example, the sequential nature of the game. Voting does not take place in an institutional vacuum. Rules of order exist, which determine the type of proposals or amendments that can be made, and the agenda of the voting process that must be used. Moreover, the voting body may use a committee structure, in which committees—or subsets of voters—discuss and reformulate proposals before they are put to a final vote on the floor. Not only the vote as such, but also these structures determine the extent to which individual players are able to affect the outcome of a vote.

The SPI refers to the ability of a player to make a difference in the outcome of a policy game. This index has many desirable features. First, it can be based on a careful and detailed analysis of some decision-making process in which the preferences of all players and all relevant institutional complexities are taken into account. Second, like traditional voting power indices, the strategic power index measures a priori power. However, in contrast to these indices the strategic power index provides a unified method to study the composite edifice of a priori and a posteriori power as a whole.

Acknowledgements We thank Ronald Heiner, Stefan Klößner, Birgit Will and a referee for helpful comments and suggestions.

References

Banzhaf, J. F. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review, 19*, 317–343.
Barry, B. (1980). Is it better to be powerful or lucky. Part I and Part II. *Political Studies, 18*, 183–194. 338–352.
Binmore, K. (1992). *Fun and games. A text on game theory*. D.C. Heath and Company: Lexington.
Braham, M., & Holler, M. (2005). The impossibility of a preference-based power index. *Journal of Theoretical Politics, 17*(1), 137–157.
Buchanan, J. (1990). The domain of constitutional economics. *Constitutional Political Economy, 1*(1), 1–18.
Felsenthal, D., & Machover, M. (1998). *The measurement of voting power*. Cheltenham: Edward Elgar.
Felsenthal, D., & Machover, M. (2001). Myths and meaning of voting power: Comments on a symposium. *Journal of Theoretical Politics, 13*, 81–97.
Garrett, G., & Tsebelis, G. (1997). Why power indices cannot explain decision-making in the European Union. In D. Schmidtchen & R. Cooter (Eds.), *Constitutional law and economics of the European Union* (pp. 11–31). Cheltenham: Edward Elgar.
Garrett, G., & Tsebelis, G. (1999a). Why resist the temptation to apply power indices to the European Union. *Journal of Theoretical Politics, 11*(3), 291–308.
Garrett, G., & Tsebelis, G. (1999b). More reasons to resist the temptation of power indices in the European Union. *Journal of Theoretical Politics, 11*(3), 331–338.
Holler, M. J., & Owen, G. (2001). *Power indices and coalition formation*. Dordrecht: Kluwer.
Laruelle, A., & Valenciano, F. (2008). *Voting and collective decision-making*. Cambridge: Cambridge University Press.
Laruelle, A., Martinez, R., & Valenciano, F. (2006). Success versus decisiveness: Conceptual discussion and case study. *Journal of Theoretical Politics, 18*, 185–205.
Machover, M. (2002). Comment on strategic power in policy games. In M. Holler, H. Kliemt, D. Schmidtchen, & M. E. Streit (Eds.), *Power and fairness* (pp. 225–227). Mohr Siebeck: Tübingen.
Napel, S., & Widgren, M. (2001). Inferior players in simple games. *International Journal of Game Theory, 30*(2), 209–220.
Napel, S., & Widgren, M. (2002). *Strategic power revisited*. CESifo Working Paper, No. 736 (2).
Napel, S., & Widgren, M. (2004). Power measurement as sensitivity analysis. A unified approach. *Journal of Theoretical Politics, 16*(4), 517–538.
Ordeshook, P. (1992). *A political theory primer*. New York: Routledge.

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

Popper, R. (1960). *The poverty of historicism*. London: Routledge/Kegan Paul.

Romer, T., & Rosenthal, H. (1978). Political resource allocation. *Controlled Agendas, and the Status Quo, Public Choice, 33*, 27–43.

Romer, T., & Rosenthal, H. (1979). Bureaucrats versus voters: On the political economy of resource allocation by direct democracy. *Quarterly Journal of Economics, 93*, 561–587.

Schmidtchen, D., & Steunenberg, B. (2002). Strategic power in policy games. In M. Holler, H. Kliemt, D. Schmidtchen, & M. E. Streit (Eds.), *Power and fairness* (pp. 203–224). Tübingen: Mohr Siebeck.

Steunenberg, B., Koboldt, C., & Schmidtchen, D. (1996). Policy making, comitology, and the balance of power in the European Union. *International Review of Law and Economics, 16*, 329–344.

Steunenberg, B., Schmidtchen, D., & Koboldt, C. (1997). Beyond comitology: European policy-making with parliamentary involvement. *Aussenwirtschaft, 52*, 87–112.

Steunenberg, B., Schmidtchen, D., & Koboldt, C. (1999). Strategic power in the European Union: Evaluating the distribution of power in policy games. *Journal of Theoretical Politics, 11*(3), 339–366.

Straffin, P. D. (1994). Power and stability in politics. In R. J. Aumann & S. Hart (Eds.), *Handbook of game theory with economic applications* (Vol. 2, pp. 1127–1151). Amsterdam: Elsevier.

Widgren, M., & Napel, S. (2001). *The power of a spatially inferior player*. The Research Institute of the Finnish Economy, Discussion Paper 771.

Part V
Political Competition and Voting Procedures

The Shapley–Owen Value and the Strength of Small Winsets: Predicting Central Tendencies and Degree of Dispersion in the Outcomes of Majority Rule Decision-Making

Scott L. Feld, Joseph Godfrey, and Bernard Grofman

1 Introduction

There are many different models of pivotal voting power that have been proposed. Most of these fall into the category of what are called *a priori* power scores. These are ones where some distribution of feasible outcomes is assumed and the probability of a given voter being pivotal is calculated wrt to that sample space based on that voter's (relative) weight in some particular voting game. One of the least known, but potentially most important of the power measures that are not *a priori* is the *Shapley–Owen value* (Shapley and Owen 1989), which is based on a uniform distribution of alternatives over a two dimensional (or higher) issue space, with the voters taken to be points embedded in that issue space, and with voter preferences customarily, for simplicity, taken to be Euclidean. The Shapley–Owen value is not regarded as an *a priori* power score since the power score (SOV) assigned to voters

The third-named author's contribution to this research was supported by the Jack W. Peltason (Bren Foundation) Endowed Chair at UCI and under Social Sciences and Humanities Research Council of Canada (SSHRCC) Grant #410-2007-2153 to study "Political Competition" (co-PIs: Stanley Winer, and J. Stephen Ferris).

S.L. Feld
Department of Sociology, Purdue University, West Lafayette, IN 47907, USA
e-mail: sfeld@purdue.edu

J. Godfrey
WinSet Group, LLC, 6107 Arlington Blvd #E, Falls Church, VA 22044, USA
e-mail: jgodfrey@winset.com

B. Grofman (✉)
Department of Political Science, Institute for Mathematical Behavioral Sciences, Center for the Study of Democracy, University of California, Irvine, CA 92697-5100, USA
e-mail: bgtravel@uci.edu

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare, DOI 10.1007/978-3-319-05158-1_16,

is a function of exactly where in the issue space these voters are located, not simply on voter weights in the voting game. In this essay we will draw on insights from the SOV in simple majority rule spatial voting games in two dimensions where voters have Euclidean preferences.

There is a long history of inquiry into the stability and predictability of majority rule processes in contexts where alternatives can be taken as points in a multidimensional issue or policy space (e.g., Plott 1967; Kramer 1972; Shepsle and Weingast 1981; Feld and Grofman 1987; Miller et al. 1989; Koehler 2001). It is widely understood that there are generally no equilibria, and that there are usually majority preferred paths that can lead from any position to any other position in the issue space (Mckelvey 1979, 1976). As Bianco et al. (2006, 2008) note, the findings of this earlier literature have been widely interpreted to imply that one can neither expect stability nor predictability of outcomes in spatial voting situations. Nevertheless, there are also incontrovertible empirical findings from experimental committee voting games that committee voting processes do reach stopping points that are not merely random. And, when we look at real world data in situations where we can estimate the ideological location of both voters and observed outcomes, e.g., wrt to voting processes such as those in the U.S. Congress or the U.S. Supreme Court, we again find a far from random pattern of outcomes relative to the distribution of estimated legislator voter ideal points,

Like Bianco et al. (2004, 2006, 2008), Schofield (1993, 1995a, b, 1999) and earlier work such as Ferejohn et al. (1984), we suggest that, even there is no core to the voting game, while all outcomes may be possible, some are more likely than others. In particular, as we shall see, the Shapley–Owen value, and insights derived from it about the underlying geometric structure of majority rule preferences, can aid us in identifying where outcomes of majority rule spatial voting games are most likely to be found.

Bianco et al. (2004, 2006, 2008) focus on the set of points in the *uncovered set* (Miller 1980, 1983) as the likely outcomes of majority rule voting processes over a "king of the hill" type agenda.[1] The *uncovered set* is the set of points such that no alternative in the set has another alternative that is both majority preferred to it and majority preferred to all point that it defeats. Another way of defining the uncovered set is as the set of points that beat all other points either directly or at one remove.[2] Thanks to new developments in computer software (Bianco et al. 2004;

[1]In a "king of the hill" agenda, there is a prevailing alternative and in pairwise fashion, some new alternative (proposal) is matched against the present "king of the hill." If the new alternative fails to receive a majority against the present king of the hill, then the process continues with a second, third, etc. alternative being proposed. If the new alternative defeats the present king of the hill, then it becomes the new king of the hill, and the process continues Either the agenda for this process is finite, e.g., a given status quo which enters the last vote, or there is some procedure for invoking cloture, so that voters can stop the process once they find an "acceptable" king of the hill.

[2]The work of Schofield we have previously cited uses a solution concept called the *heart*, which seems very appropriate for weighted voting coalition games with a limited number of players, such as multiparty cabinet formation games, where ideal points are to a large extent a matter of common

Godfrey 2007) it is now possible to identify the location of solution concepts such as the uncovered set even for games with large numbers of actors, even though an analytic solution for the uncovered set is known only for the three-voter case (Feld et al. 1987; Hartley and Kilgour 1987).

Looking at results over a 20 year period of experimental research, Bianco et al. (2006) show that around 90 % of all the observed outcomes in nearly a dozen five person experimental committee voting games lie within the uncovered set. The Bianco et al. (2006) article represents, in our view, a major theoretical breakthrough in that, until their work, except in games where there was a core (where the prediction that outcomes in experimental (committee voting) games would tend toward the core was strongly supported), there simply was not a satisfactory game theoretic model to predict where outcomes would lie in committee voting games. The absence of satisfactory theory for non-core situations is highlighted in the discussion of results in Fiorina and Plott (1978) for their non-core game, and similar language is found in later experimental work on committee voting games up until very recently. Moreover, in our view, committee voting experiments trailed off after the late 1980s in part because of the absence of reliable theory that could be tested and further extended, while experimental work focused on areas, such as the study of auctions, where theory with real predictive bite was much better established.

Building on the Bianco et al. (2006) work on the predictive power of the uncovered set, we take a different, albeit related, tack. We will look for mechanisms that can explain why outcomes of committee voting games are likely to be in the uncovered set. This exploration will take us away from the uncovered set, *per se*, to look, instead, for even more general features of the structure of majority rule in the spatial voting context, features that we will demonstrate to be directly linked to the Shapley–Owen value.

The "winset" of a point is the set of other points that a majority of voters prefer to that point. Saying that there are no equilibria is equivalent to saying that all points have non-empty winsets. Nevertheless, the sizes of those winsets can vary widely. The simple intuition we propose is that, at least for king of the hill type agendas (and probably far more broadly) the size of a point's winset is a major determinant of whether a point is likely to be proposed, whether it is likely to be majority adopted, and whether is likely to be a stopping point of the voting process.

First, when a point is proposed, points with smaller winsets are more likely to be adopted because, by definition, points with smaller winsets are majority preferred to more possible status quos than other points. Second, points with smaller winsets are more likely to become the stopping point of the voting game because a majority is likely to recognize that it is difficult and unlikely for them to find and adopt a position that would be better for them, because points with small winsets will, by definition, offer few such alternatives that can defeat them and so, in a king of the

knowledge. Schofield (1999) shows that the uncovered set is a subset of the heart. In this essay we focus on committee voting games rather than coalition games, and we will draw our comparisons to the uncovered set rather than the heart.

hill type agenda, proposals to replace them are likely to fail (or at least to require a time consuming search).

This line of argument gives rise to two very straightforward hypotheses about majority rule processes.

Empirical Hypothesis 1 Outcomes of a majority rule process are more likely to be points with smaller winsets than points with larger winsets.

Empirical Hypothesis 2 Outcomes of a majority rule process tend to center around the point with the smallest winset.

Understanding the practical implications of these hypotheses for majority rule voting games requires us to draw on theoretical insights from Shapley and Owen (1989) about the Shapley–Owen value. In particular.

Theoretical Proposition 1 (Shapley and Owen 1989) For Euclidean majority rule voting games in two dimensions, the point with the smallest winset, referred to by Shapley and Owen (1989) as the *strong point*, is located at the weighted average of the voter ideal points in the game, where the weights are simply each voter's Shapley–Owen value, i.e., the proportion of median lines on which each voter is pivotal.

The *strong point* is the spatial analogue of the *Copeland winner* in finite alternative games, i.e., the point that is defeated by the fewest other points (Straffin and Philip 1980).

Theoretical Proposition 2 (Shapley and Owen 1989) For Euclidean majority rule voting games in two dimensions, for alternatives located along any ray from the strong point, the size of winsets increases with distance from the strong point. Even more specifically, the winset of any point has an area equal to the area of the winset of the strong point plus pi times its squared distance from the strong point.

Corollary to Theoretical Proposition 2 The larger the winset of the strong point itself, the less the *relative* difference in winset size as the distance to the strong point increases.

From this theoretical result about differences in win-set size as we move away from the strong point tied to the size of the strong point's winset, we are led to our third empirical hypothesis—one that allows for a prediction about comparisons of results across different experimental voting as a function of the location of the voter ideal points in those games and the concomitant size and win set area of the strong point.

Empirical Hypothesis 3 The smaller is the winset of the strong point itself, the closer, *ceteris paribus*, will be the outcomes of a majority rule process to the strong point, and the lower the variance of the observed outcomes.

In particular, in the limit, when the strong point shrinks to a single point, the core, with an empty winset, we expect outcomes to be very close to this core—

a result which conforms to what has previously been found in studies involving experimental committee voting games with a core.

In the next section: (1) We provide some illustrative examples of winsets for majority rule processes in two-dimensional spatial contexts. (2) We formally describe the majority rule processes that have been used in experiments, and show the geometry of some of the spatial voting games used in these experiments. (3) We show that analyses of outcomes of these experiments are consistent with our theoretical predications. (4) We analyze not just final outcomes but also intermediate proposals in a few of these experiments to illustrate the plausibility of our proposed links between winset size and final outcomes. In the concluding discussion, after summarizing our empirical findings, we consider how our theory helps to explain the prediction success of the "uncovered set" as a solution concept.

2 Theoretical Properties of Winsets and Empirical Results About the Predictive Power of Winset Size for Outcomes in Experimental Games

2.1 Winsets in Majority Rule Processes in Two Dimensions

We begin consideration of winsets in spatial voting situations with a simple example. Suppose that there is a group of faculty deciding on the requirements for their graduate program in the context of a two-dimensional space, where the horizontal dimension is the number of requirements, and the vertical dimension is the relative emphasis on qualitative versus quantitative research approaches. For the purposes of illustration, Fig. 1a shows the current set of graduate program requirements as the origin in the graph. Suppose that there are three voters: "quant" who prefers more extensive quantitative requirements with an ideal point to the upper right; "qual" who prefers somewhat more extensive qualitative requirements with an ideal point to the lower right; and "easy" who just prefers less extensive requirements than currently in place, with an ideal point somewhat to the left. The status quo point in this example has a winset as shown in the Fig. 1a.

As noted earlier, Shapley and Owen (1989) show that the point with the winset of minimum area, called the *strong point*, is a weighted average of the locations of all the voter ideal points, where the weights are determined by the ranges over which each ideal point determines the boundaries of the winset, i.e., the range of angles over which the voter is pivotal. When there are only three voter ideal points, the angles of the triangle connecting those points turn out to be the relevant weights. Thus, the voter who subtends the largest angle has the single greatest influence on the location of the strong point. For the example presented above, the relevant angles are highlighted in Fig. 2 below.

It can be seen that qual has the largest angle; easy has the next; and quant has the smallest in this situation. The point with the smallest winset is the locations

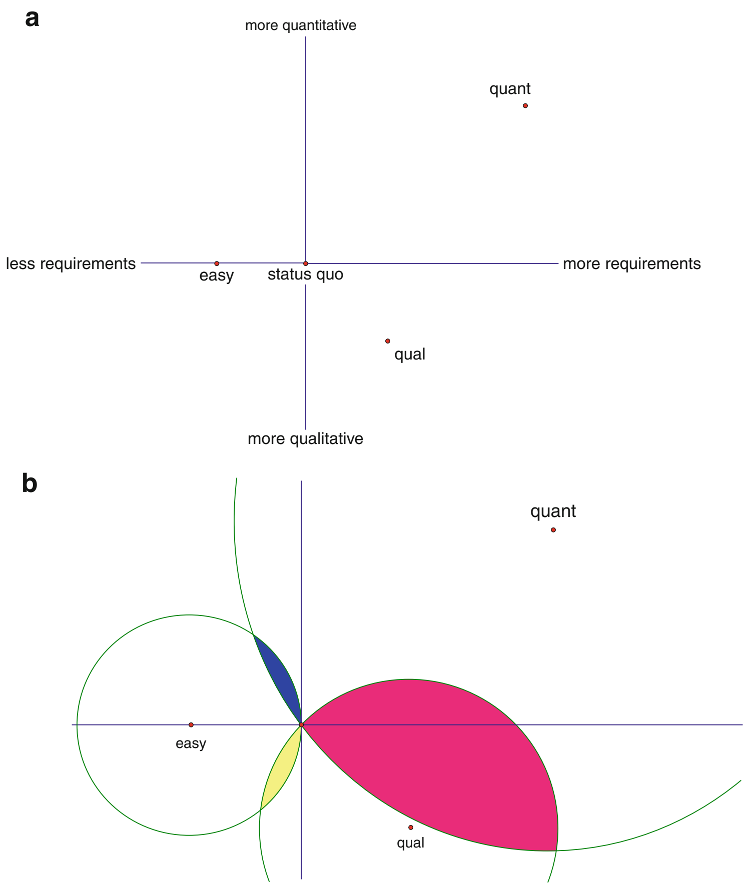
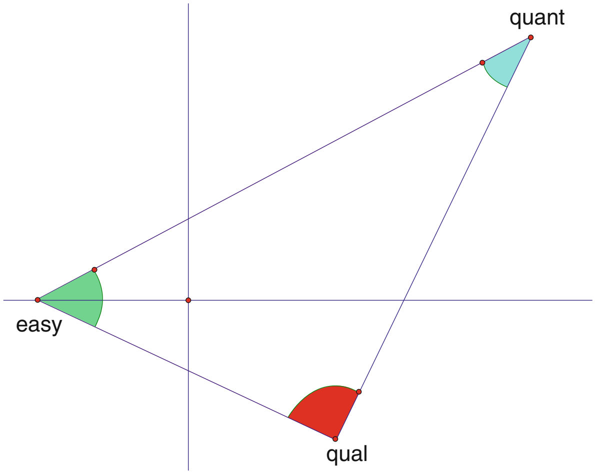


Fig. 1 A three voter game and the Winset of a status quo point in the game. (**a**) A three voter example of a committee voting game. (**b**) The Winset of the status quo in (**a**)

of the voter ideal points weighted by their angles. In this case the weights turn out to be approximately 0.5, 0.3, and 0.2, respectively. That weight-averaged point "strong point" is shown with its winset in Fig. 3a. As noted earlier, Shapley and Owen (1989) further prove that the winsets of points increase in size directly as the squared distance from the strong point. Thus, points that are equidistant from the strong point have equal size winsets. The circles in Fig. 3b indicate sets of points with equal size winsets.

Recall that our theoretical prediction is that points with smaller winsets are more likely to be the endpoints of sequential majority rule voting processes. Therefore, our theory implies that points that are closer to the strong point are more likely to be outcomes of majority rule processes than points further away, *ceteris paribus*. Furthermore, since winset sizes increase equally in all directions from the strong

Fig. 2 Angles used as weights to determine the location of the strong point in the three-voter example of Fig. 1



point, outcomes are equally likely to be in any direction from the strong point, and are therefore expected to center around the strong point.[3]

In these games the process of voting is carried out by independent individuals with incomplete information. In particular, unlike the situation with voting games over a finite set of alternatives, sophisticated voting in the sense of Farquharson (1969), which requires the ability to use backward induction to identify sophisticated strategies (Mckelvey and Niemi 1986), is simply not possible. Furthermore, in the usual setting of experimental games, where information about voter ideal points is withheld from the players, individuals cannot calculate the sizes of winsets and, in any case, have no particular personal interest specifically in supporting points with smaller winsets. Their interests are in supporting points that are closer to their ideal points. Any strategy they employ that goes beyond that is likely to be highly variable from individual to individual. Nevertheless, we suggest that the effect of the "social" process driven by the preferences of the voters for outcomes closer to themselves is likely to result in overall outcomes being points with smaller winsets. The key intuition is that points with small(er) win-sets, once chosen, are "hard(er)" to defeat.

At the same time, we recognize that there are many points with very similarly sized winsets. For example, points near the strong point have winsets that are only very slightly larger than the winset of the strong point. It is unlikely that any process that is driven by the relative sizes of winsets can make fine distinctions. Thus, points very near to the strong point are essentially equally likely to be the end of the process

[3]Note that points outside of the Pareto Set are unlikely to be outcomes of these types of processes even when they have relatively small winsets, and the strong point can sometimes be near the boundary of the Pareto Set. Consequently, we expect that situations where the strong point is close to the boundary of the Pareto Set will be exceptions to our general expectation that the strong point will be central among the outcomes.

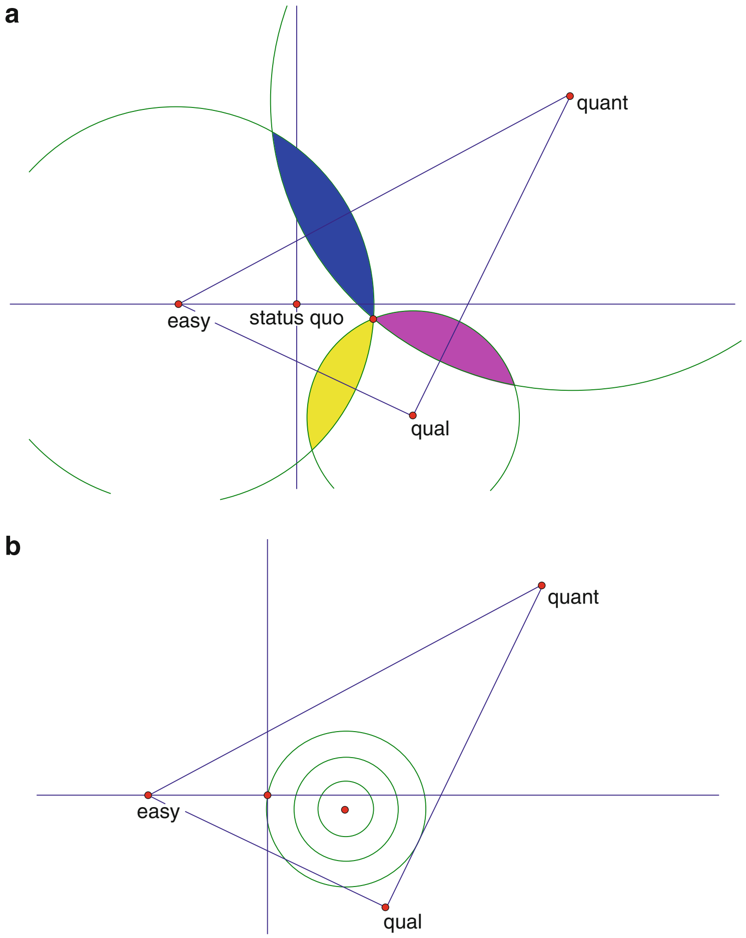


Fig. 3 The alternative with smallest Winset, and Winset sizes as we move away from that point in the three voter example of Fig. 1. (**a**) The Winset of the strong point. (**b**) Circles with equal size Winsets around the strong point

as the strong point itself. Also, as winset size increases with distance from the strong point, there will also be increasing numbers of points with winsets of each larger size**;** specifically, each concentric circle around the strong point can be thought an iso-winset line, and the further out circles are larger and contain more points. Thus, as we move further from the strong point, there are more points to choose from even though each is chosen with a lower probability.

It is impossible to specify a general functions describing the expected distance of outcomes from the strong point, based upon the countervailing effects of the increasing availability of points and the declining likelihood of each particular point, because the specific likelihood of outcomes at each distance will depend upon many factors determining the declining marginal utility to the particular actors, their ability to detect such differences under the particular circumstances of play, and whatever might influence their willingness and abilities to do anything about it, among other things.

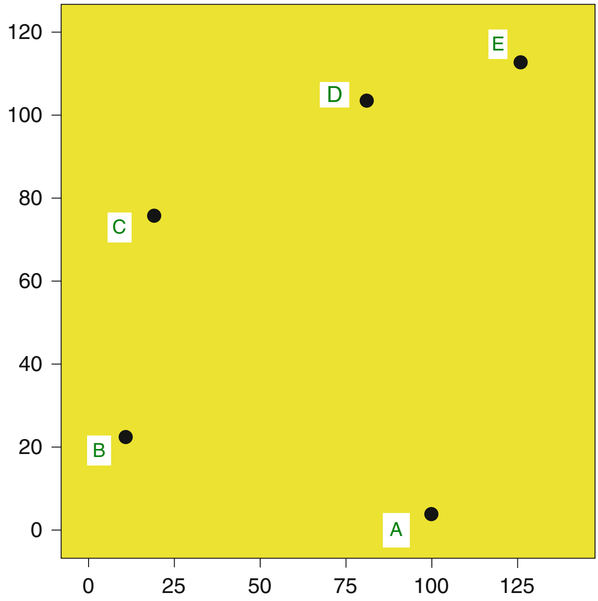
Nonetheless, we can hyopothesize with some confidence that, *ceteris paribus*, the faster that winsets in any given voting game grow with distance from the strong point, the closer that the outcomes will be, on average**,** to the strong point. We suggest that the relevant rate of growth is relative, i.e. the proportionate increase, rather than the absolute increase. The absolute size of the winsets always increase as pi d-squared, where d is the distance from the strong point. However, the relative importance of those increases declines with the size of the smallest winset, the winset of the strong point itself. This leads us to our third theoretical prediction, namely that the variation among the outcomes around the strong point will increase with the size of the winset of the strong point. When the winset of the strong point is very small, then the outcomes are likely to cluster relatively closely to the strong point. However, when the winset of the strong point is large, then the variation of the outcomes will be larger.

The variance around the strong point will depend, however, not only upon the size of the winset of the strong point, but also upon other factors previously suggested**,** e.g., the many factors determining the declining marginal utility to the particular actors, their ability to detect such differences under the particular circumstances of play, and whatever might influence their willingness and abilities to do anything about it, among other things. For example, we might expect that anything that makes players more nervous or impatient will lead them to be more willing to accept and vote to end at outcomes that they would not otherwise accept—that would imply greater variation in outcomes overall. But there is a strong *ceteris paribus* operating in our analyses: we simply do not know enough about how the selection of players, experimental instructions, and the play of the game itself may affect variation of outcomes around the strong point**.**

2.2 Majority Rule Committee Voting Experiments in Spatial Contexts

Bianco et al. (2006, 2008) review most of the experiments conducted by a variety of different researchers on majority rule processes in a spatial context that involve committee decision making. Fiorina and Plott published their classic experiment in 1978, and established the paradigm for subsequent experiments. The typical procedures in these experiments have involved five subjects voting for a point on a two-dimensional map. A session begins with a status quo point determined by the researcher. By various procedures, a proposal for an alternative positions in the space. Is proposed. Then, the group votes on whether or not to replace the current status quo with the proposed alternative. If a majority of the voters prefer

Fig. 4 The "Skew Star" five voter experimental game



the alternative, then that alternative becomes the new status quo point. Then, a new alternative is proposed and voted upon, etc. The process ends when a voter proposes stopping, and a majority of the voters approve of stopping at that point. Researchers have modified these procedures and limited the alternatives that can be proposed in various ways for various theoretical purposes, but the basic procedures have been similar in several experiments.

Figure 4 shows a typical situation, this one drawn from experiments conducted by Laing and Olmstead (1978). They called this their "Skew Star" situation.

In the game shown in Fig. 4, a hypothetical sequence of votes might move the status quo around the space as shown in the following hypothetical example (see Fig. 5).

Just as with the simple three voter situation in the previous section of this paper, each possible status quo point in this five voter game has a winset that consists of a set of petals, where each petal is the set of points that are preferred to the status quo point by some majority of the voters. Some points have smaller winsets than others as shown in Fig. 6. For example, point S has a smaller winset than T.

There is a single point with the smallest winset, and winset size increases with the squared distance to that strong point. The strong point and its winset is shown in Fig. 7.

The actual set of outcomes for the 18 experimental runs conducted by Laing and Olmstead (1978) for this game are shown in Fig. 8, which also shows, for comparison purposes, the mean location of the outcomes in the game as well as the location of the strong point.

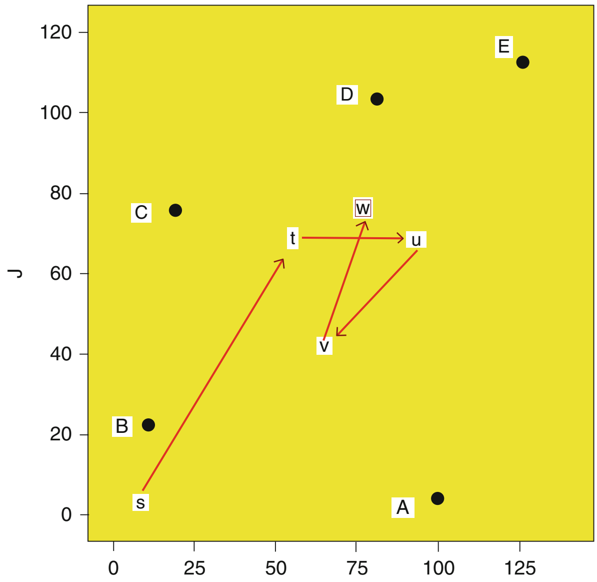


Fig. 5 A hypothetical trajectory of votes in the Skew Star game of Fig. 4

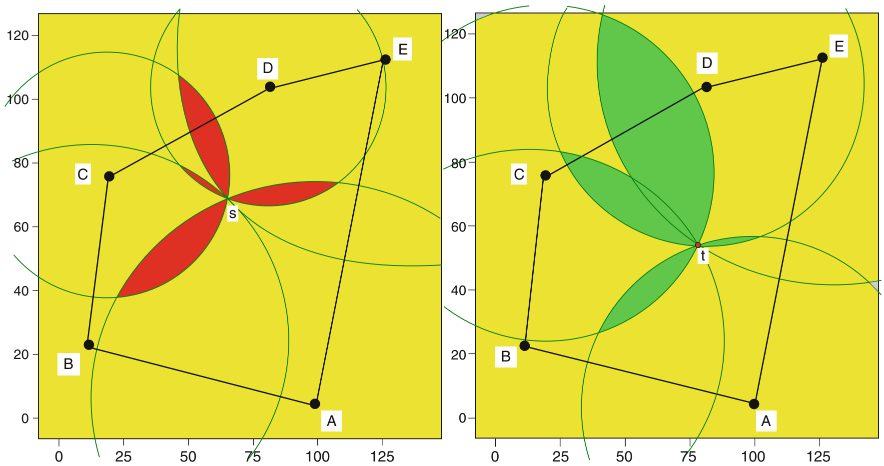


Fig. 6 Winsets of two points in the Skew Star game of Fig. 4

As noted previously, we have proposed three empirical hypotheses about the outcomes of spatial voting games, which we may summarize as below:

1. Points with smaller winsets are more likely outcomes than points with larger winsets.
2. The outcomes will tend to center on the strong point.

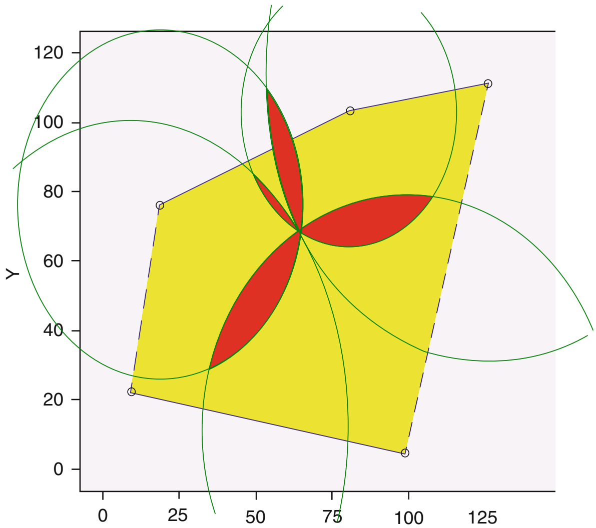


Fig. 7 The Winset of the strong point in the Skew Star game shown in Fig. 4

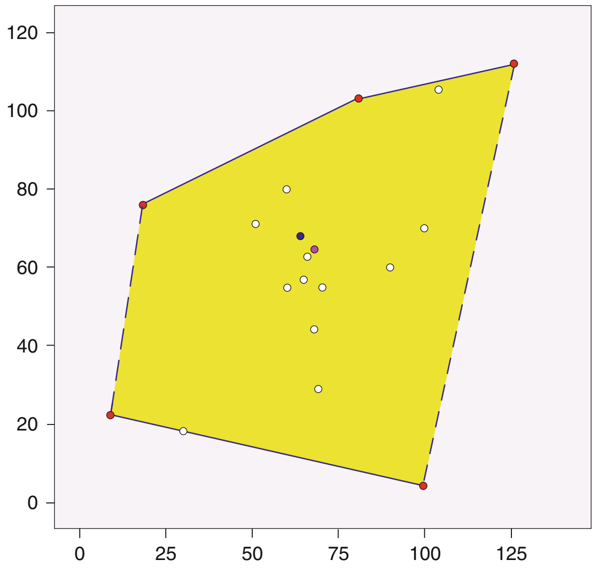


Fig. 8 The experimental outcomes in the Skew Star game shown in Fig. 4, showing the mean location of the outcomes in *black* and the location of the strong point in *red* (color figure online)

3. The variance of outcomes from the strong point will be smaller the smaller is the size of the winset of the strong point.

There is support for both of our hypotheses (Figure omitted for reasons of space). First, as expected, the outcomes are disproportionately clustered near to the strong point, i.e., are among the points with smallest winsets.

Second, also as expected, the strong point is located relatively centrally among the experimental outcomes because winset sizes increase symmetrically around the strong point with distance from the strong point. Nevertheless, the outcomes include some that are fairly far away from the strong point.

Our theory (Hypothesis 3) also suggests that there will be greater variation in ourtcomes when the size of the winset of the strong point itself is relatively large. In the Skew Star game of Fig. 4, as we will see when we present comparisons of this game to other games later in the paper, the win set of the strong point in the game is relatively large with respect to the Pareto set (see Fig. 8), and so winset sizes will rise only slowly with distance from the strong point, and thus, as expected, we get a fairly considerable scatter of outcomes around the strong point (see Fig. 9).

These detailed data from this one experiment are presented merely to illustrate how we use experimental findings to examine and test the implications of our hypotheses. The data for all the relevant experiments are analyzed more systematically in the next section.

2.3 Testing Our Hypotheses Using a Large Body of Data on Experimental Outcomes in Committee Voting Games

Using data from the same experiments reanalyzed by Bianco et al. (2006) and additional experiments that Bianco et al. (2008) conducted themselves, we reanalyze games used in 18 different experiments by seven different teams of researchers. Two of these games were initially used by Mckelvey and Ordeshook under several different experimental conditions and then used again by Endersby (1993) under other experimental conditions. For our present purposes, we combined all the data collected for the same games even if conducted under different conditions by different experimenters. Thus, we were able to reanalyze the results from a total of ten different games.

The experiments whose outcomes are used here were conducted for a variety of different specific purposes, including testing different solution concepts under somewhat different structural conditions. For the present purposes we are ignoring the relatively small differences in experimental procedures among the experiments to focus on the overall tendencies that emerge even when there are some potentially confounding differences among the experimental protocols.

First, we find that the mean positions among the experimental outcomes in all of these games are very close to the strong points of the games. Table 1 shows the coordinates of the mean outcome compared with the coordinates of the strong point

Table 1 Mean outcomes as compared to the strong point in the game

Game	Mean outcome	Strong point	Pareto set area
Bianco 1	(38,73)	(41,78)	2450
Bianco 2	(67,20)	(67,19)	1673
Fiorina_Plott_1978	(45,63)	(47,62)	4647
Laing_Olmsted_1978_A2_The_Bear	(83,57)	(80,55)	7875
Laing_Olmsted_1978_B_Two_Insiders	(66,34)	(61,35)	5106
Laing_Olmsted_1978_C1_The_House	(76,55)	(90,53)	9311
Laing_Olmsted_1978_C2_Skew_Star	(69,67)	(64,69)	8190
McKelvey_Ordeshook_Winer_1978	(84,122)	(88,115)	10845
PH	(57,36)	(58,36)	2753
PHR	(67,34)	(70,32)	2763

along with the area of the Pareto set so that we can see how close the strong point is to the mean outcome in each game relative to the size of the Pareto set.[4]

We see from Table 1 that mean outcomes are close to the strong point in nine of ten games, and close in one dimension, but not so close in the second dimension, in the remaining game. For each of the ten games, we did significance tests to determine whether the mean for the x coordinate was statistically significantly different from the x coordinate of the strong point, and similarly for the y-coordinate. For these 20 significance tests, two of them were statistically significant, which is close to what would be expected by chance alone ($p = 0.05$) if the means for the populations were exactly at the strong points.

Second, not only do the outcomes tend to tend to be close to the strong point, on average, but they are also close to the strong point when we think of distance in terms of the size of the Pareto set. In general, the distance between the means of the outcomes and the strong points are less than 2 % of the sizes of the Pareto sets.

Third, outcome variance tends to be related to the size of the winset of the strong point, as we theoretically predicted. When the winset of the strong point is smaller, relative to the size of the Pareto set, then there is less variation in the outcomes around the strong point (again relative to the size of the Pareto set), as is shown in Table 2.

Over this small set of ten games the correlation between outcome variance in the game and the size of the winset of the strong point in the game is $+0.21$. It would be much stronger except for two outliers. The Bianco two game has a strong point with

[4]Consider, for example, the strong point (shown in black) in Fig. 10. It is very close to the mean location of the outcomes (shown in pink) when we think of closeness relative to the spread of the voter ideal points. The Pareto set in these situations is the convex figure that is enclosed by all the lines between the voter ideal points. For any point outside the Pareto set, the voters always unanimously prefer some other point inside the Pareto set. Consequently, voters generally have little reason to ever propose alternatives outside of the Pareto set, and they rarely do so. Consequently, the effectiveness of prediction should be considered with respect to proposals in the Pareto set (shaded yellow in Fig. 10).

Table 2 Winset size and outcome variance

Game	Winset size	Variance in outcomes
Fiorina_Plott_1978	0.02	0.02
Bianco_2_2003	0.02	0.14
McKelvey_Ordeshook_Endersby_PHR	0.06	0.03
McKelvey_Ordeshook_Winer_1978	0.06	0.05
Laing_Olmsted_1978_C1_The House	0.06	0.15
McKelvey_Ordeshook_Endersby_PH	0.07	0.04
Laing_Olmsted_1978_C1_The Bear	0.11	0.06
Laing_Olmsted_1978_C1_Skew Star	0.11	0.11
Bianco_1_2003	0.12	0.1
Laing_Olmsted_1978_C1_Two_Insiders	0.15	0.11

a small winset, but has considerable variance in outcomes, and the Laing-Olmsted House game also has considerably more variation than would have been expected. If these two outliers are omitted, the correlation for the remaining eight games is 0.90.

Further examination of these two games with unexpectedly high variation in outcomes indicates that the high variation in each case arises from just a couple of extreme outlying outcomes. A closer examination of the data from the Bianco two game (where we were able to examine the whole process for each experimental run) indicates that the two extreme outcomes in that game occurred when the voters made the rare decision to stop immediately after accepting their first proposal. Such findings suggest that there may be idiosyncratic noise in any play of any particular play of a game with a small number of players. However, the rest of the pattern suggests that apart from such "noise", the outcomes are consistent with our theoretical expectations that there is generally less variance in outcomes when the strong point has a smaller winset.

2.4 Proposing, Adopting, and Stopping in Experimental Committee Voting Games

Unfortunately, the data on proposed and adopted points are not included in the published reports on any of the experimental studies used in the previous analyses. However, William Bianco and his colleagues (personal communication, 2009) have generously provided us with these data from their recent experiments with two particular experimental games. These data provide a relatively small number of cases, but are sufficient to provide some preliminary findings. We combine the results of the two games in Table 3 below. The starting point (also provided to us by Professor Bianco) is not included in the analyses shown in that table.

In both of these games, it is clear that the winset sizes of proposed points are considerably smaller than for other Pareto points, that winsets of adopted points

Table 3 Mean and standard deviations of Winset sizes for different sets of points

	Bianco game 1			Bianco game 2		
	Mean	SD	*n*	Mean	SD	*n*
Pareto points	1314	1463		2210	2118	
Proposed points	531	882	261	785	1499	185
Adopted points	386	396	147	446	770	91
Stopping points	246	236	28	244	641	28

are considerably smaller than other proposed points, and that stopping points have considerably smaller winsets than other adopted points. Even with these relatively small sample sizes, all of these differences are statistically significant (with $p \le 0.05$ using 1-tail tests). Thus, as hypothesized, points with smaller winsets are more likely to be majority approved, and chosen as stopping points than other points in the Pareto. Moreover point with smaller winsets are also more likely to be proposed.[5]

3 Conclusions

3.1 Key Findings

Until Bianco et al.'s recent publications, previous research seems to have led researchers to the conclusion that there was no good theory to predict the outcomes of experimental committee voting games in two or more dimensions.[6] Bianco et al. reopened the question with their findings that nearly all outcomes in large body of experimental voting games fell within the uncovered set, and that such results were considerably more likely than would be expected by chance. Our approach to this same data has emphasized the predictive power of small win sets.

[5]It is important to recognize that, unless there is a core to the voting game, it need not be true that points that beat other points have smaller winsets than the points they beat. At the start of the process, when the points are relatively far out from the strong point, there is a tendency for the process to move inward. However, once the status quo is further in toward the strong point, there is no necessary expectation that further points will have smaller winsets. In fact, nearly all the points in the winsets of points close to the strong point have winsets larger than the strong point itself—consequently, if the process does not stop at the strong point, it necessarily moves to points with a larger winset than the strong point itself. Nonetheless, if the outcome is a point near the strong point this will be a point with a relatively small winset.

[6]On the other hand, there are models of spatially embedded coalition formation games and of party competition games that do generate empirically testable models that garnered considerable empirical support. Trying to reconcile the theoretical and empirical findings on committee voting, coalition formation, and party competition, however, takes us into issues well beyond the scope of this paper.

While predicting that points with smaller winsets are more likely outcomes does not provide any specific boundaries for the set of predicted outcomes, it does allow us to make some specific predictions. First, we predicted that points with smaller winsets are more likely outcomes than points with larger winsets. Second, since winset size is distributed symmetrically around the point with smallest winset, the strong point, we predicted that outcomes will center on the strong point. Third, since winset size increases as a specific monotonic function of distance from the strong point, and consequently relative winset size increases more slowly when the winset of the strong point itself is large (relative to the Pareto set), we predict that the outcomes will diverge further from the strong point in games when the winset of the strong point is large than in games when the winset of the strong point is small. Evidence from 17 experiments using 10 different experimental games confirms each of these predictions, and suggests that previous findings concerning the success of the uncovered set may result from the fact that points in the uncovered set tend to have small winsets.

Movement toward points with smaller winsets can be considered as a "centrifugal" force pulling outcomes toward the strong point. However, of course, we recognize that there are centripetal forces that may pull outcomes somewhat away from the strong point. For example, actors may tend to make proposals for alternatives that are close to their ideal points, and voters may accept outcomes that are good enough, even if not ideal. Also, any (minimal) winning coalition can exert total control of outcomes, and such coalitions may pull outcomes toward the hull of that coalition, which might not include the strong point. Furthermore, there may be confusions or misperceptions that also affect outcomes, and some voters may be more attentive to the voting process than others. Each of these aspects of the game (e.g., satisficing, coalition formation processes, variation in information levels or actor involvement) can pull outcomes away from the strong point. Moreover the specific voting rules (e.g., whether a defeated alternative can be reconsidered) and other features of the experiment (e.g., how much knowledge each voter has about the preferences of the other voters)[7] may matter a great deal, suggesting the desirability of additional experiments for a fixed set of voter locations to see how rules of the game and other context features matter for the mean and variance of outcomes and for speed of convergence. Nevertheless, we believe that size of win sets provides such a strong gravitational pull on outcomes that it will serve as a key theoretical tool for understanding and predicting outcomes and outcome trajectories not just in king of the hill spatial committee voting games, but also in a wider set of committee voting games, and in real world politics that can be modeled as voting over multidimensional issues.

[7] To the extent that voters can develop a sense of the preferences of other players, points perceived as "more fair" may be more likely to be proposed and accepted as the final outcome, or perhaps, points that are perceived to be likely to defeat other alternatives, e.g., points on the boundary of a minimum winning coalition, may be more likely to be proposed (cf. the notion of the *competitive solution* in Mckelvey et al. 1978).

Table 4 Mean squared distance to the strong point from the Pareto, the uncovered set and the outcomes in the game that lie in the uncovered set

	Mean D squared		
Game	Pareto	UC	Outcome in UC
Bianco 1	1314	312	225
Bianco 2	2210	50	21
Fiorina_Plott_1978	2360	128	64
Laing_Olmsted_1978_A2_The_Bear	1569	853	323
Laing_Olmsted_1978_B_Two_Insiders	1304	654	454
Laing_Olmsted_1978_C1_The_House	2059	939	500
Laing_Olmsted_1978_C2_Skew_Star	1659	890	397
McKelvey_Ordeshook_Winer_1978	2364	1041	594
PH	601	281	93
PHR	597	290	100

3.2 Reconsidering the Success of the Uncovered Set as a Predictor of Experimental Game Outcomes

Part of the motivation for the present paper comes from recent publications by Bianco and colleagues reporting their findings that the *uncovered set* is a very successful solution concept for experimental committee voting games. While our empirical findings are only that we do as well in predicting outcomes with win-set size as we do with the uncovered set (taking our winset prediction set to be the same size as the uncovered set), we would argue that there are good reasons to prefer the winset explanation for observed experimental outcomes.

1. It is highly plausible that alternatives that defeat most other alternatives are less likely to be defeatable by a randomly chosen other alternative than points with a larger win set, and thus are more likely to end up ultimately chosen. In contrast there really is no comparably good “story” to explain the predictive success of the uncovered set.
2. Outcomes within the uncovered set have smaller winset sizes than general points in the Pareto set.[8] Table 4 shows that this is true for all of the experimental games that Bianco et al. analyzed and that we reanalyzed above. This results suggest that the correlation between being in the uncovered set and having a small winset may account for the predictive success of the uncovered set.[9]

[8]The uncovered set consists of points with small winsets because points with large winsets are likely to be covered by some other point with a smaller winset (see Miller 2007).

[9]Bianco et al. (2004, 2006, 2008) note that when the uncovered set is large, the uncovered set can include most of the points in the Pareto Set, so predictions based on the uncovered set are not that specific, though they predict far better than chance. They are equally well aware that, in the unusual situations where the uncovered set is small, e.g. when there is a core, then some observed outcomes in experimental voting will not lie exactly at the core and thus will fall outside the uncovered set.

On the other hand, we would note that the evidence we have presented for the strong point determined by Shapley–Owen values being the center of the distribution of the observed outcomes of experimental spatial majority rule voting games must be interpreted with some caution, since there are other solution concepts that are also located very centrally in the Pareto set and very close to the strong point, e.g., the centroid of the uncovered set or the center of the *yolk*, the center of the smallest circles that touches all median lines.[10] The evidence presented in this paper does not really allow us to distinguish the hypothesis that points are centered around the strong point from the hypothesis that points are centered around the center of the yolk, or the centroid of the uncovered set.[11] It is only with further experimental work, especially work that allows us to examine what points are proposed as well as what points remain "king of the hill," that we will be able to devise critical tests among competing explanatory models.

References

Bianco, W. T., Jeliazkov, I., & Sened, I. (2004). The uncovered set and the limits of legislative action. *Political Analysis, 12*, 256–276.

Bianco, W. T., Lynch, M. S., Miller, G. J., & Sened, I. (2006). A theory waiting to be used: a reanalysis of canonical experiments on majority rule decision-making. *Journal of Politics, 68*, 837–850.

Bianco, W. T., Lynch, M. S., Miller, G. J., & Sened, I. (2008). The constrained instability of majority rule. *Political Analysis, 16*, 115–137.

Endersby, J. W. (1993). Rules of method and rules of conduct: an experimental study on two types of procedure and committee behavior. *Journal of Politics, 55*, 218–236.

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

[10]Bounds on the uncovered set are often stated in terms of the center of the yolk (in two dimensions, the yolk is the smallest circle touching all median lines), e.g., in classic work, Mckelvey (1986) shows that the uncovered set must lie with 4 yolk radii of the center of the yolk, and this bound has been tightened by others (Feld et al. 1987; Miller 2007). Thus, it is natural to ask about the relationship between the center of the yolk and winset size, on the one hand, and the relationship between the center of the yolk and the strong point, on the other. Craig Tovey (personal communication, 2009), investigating a conjecture by Scott Feld, has recently proved a result closely related to the Shapley and Owen result bounding the size of winsets by the size of and distance to the strong point, namely that bounds on the size of the winset of any point can be stated in terms of the size of the winset of the strong point and the distance of the point to the center of the yolk. As a corollary, he also shows that the strong point can be no more than 2.16 yolk radii from the center of the yolk. Although that is the tightest bound known, we have not found any situations where the strong point is more than one yolk radius from the center of the yolk, and we believe it must be within the yolk in the three voter case. In the games reported on here, the yolk is considerably closer to the strong point than 2.16 yolk radii.

[11]Similarly, if voters focus their attention on one dimension at a time, resulting in an outcome at the generalized median, i.e., at the location of the respective median voter along each of two dimensions (Shepsle and Weingast 1981; cf. Feld and Grofman 1988), the generalized median cannot be very far away from the strong point.

Feld, S. L., & Grofman, B. (1987). Necessary and sufficient conditions for a majority winner in N-dimensional spatial voting games—an intuitive geometric approach. *American Journal of Political Science, 31*, 709–728.

Feld, S. L., & Grofman, B. (1988). Majority rule outcomes and the structure of debate in one-issue-at-a-time decision making. *Public Choice, 59*, 239–252.

Feld, S. L., Grofman, B., Hartley, R., Kilgour, M. O., & Miller, N. (1987). The uncovered set in spatial voting games. *Theory and Decision, 23*, 129–156.

Ferejohn, J. A., Mckelvey, R. D., & Packel, E. W. (1984). Limiting distributions for continuous state Markov voting models. *Social Choice and Welfare, 1*(1), 45–69.

Fiorina, M. P., & Plott, C. R. (1978). Committee decisions under majority rule: an experimental study. *American Political Science Review, 72*, 575–598.

Godfrey, J. (2007). *WINSET: a computer program for spatial voting*. Falls Church, VA: WinSet Group.

Hartley, R., & Kilgour, M. (1987). The geometry of the uncovered set in the three-voter spatial model. *Mathematical Social Sciences, 14*(2), 129–155.

Koehler, D. (2001). Instability and convergence under simple-majority rule: results from simulation of committee choice in two-dimensional space. *Theory and Decision, 50*(4), 305–332.

Kramer, G. (1972). Sophisticated voting over multidimensional spaces. *Journal of Mathematical Sociology, 2*, 165–180.

Laing, J. D., & Olmstead, S. (1978). An experimental and game-theoretic study of committees. In P. C. Ordeshook (Ed.), *Game theory and political science* (pp. 215–281). New York: New York University Press.

Mckelvey, R. D. (1976). Intransitivities in multidimensional voting models and some implications for agenda control. *Journal of Economic Theory, 12*, 472–482.

Mckelvey, R. D. (1979). General conditions for global intransitivities in formal voting models. *Econometrica, 47*, 1085–1112.

Mckelvey, R. D. (1986). Covering, dominance, and institution-free properties of social choice. *American Journal of Political Science, 30*(2), 283–314.

Mckelvey, R. D., & Niemi, R. G. (1986). A multistage game representation of sophisticated voting for binary procedures. *Journal of Economic Theory, 18*(1), 1–22.

Mckelvey, R. D., Winer, M. D., & Ordeshook, P. (1978). The competitive solution for N-person games without transferable utility, with an application to committee games. *American Political Science Review, 72*(June), 599–615.

Miller, N. R. (1980). A new 'solution set' for tournaments and majority voting. *American Journal of Political Science, 24*, 68–96.

Miller, N. R. (1983). The covering relation in tournaments: two corrections. *American Journal of Political Science, 27*(2), 382–385.

Miller, N. R. (2007). In search of the uncovered set. *Political Analysis, 15*(1), 21–45.

Miller, N. R., Grofman, B., & Feld, S. L. (1989). The geometry of majority rule. *Journal of Theoretical Politics, 1*(4), 379–406.

Plott, C. R. (1967). A notion of equilibrium and its possibility under majority rule. *American Economic Review, 57*, 787–806.

Schofield, N. (1993). Political competition and multiparty coalition governments. *European Journal of Political Research, 23*(1), 1–33.

Schofield, N. (1995a). Democratic stability. In J. Knight & I. Sened (Eds.), *Explaining social institutions*. Ann Arbor: University of Michigan Press.

Schofield, N. (1995b). Coalition politics – a formal model and empirical-analysis. *Journal of Theoretical Politics, 7*(3), 245–281.

Schofield, N. (1999). The heart and the uncovered set. *Journal of Economics-Zeitschrift Fur Nationalokonomie Supplement, 8*, 79–113.

Shapley, L., & Owen, G. (1989). Optimal location of candidates in ideological space. *International Journal of Game Theory, 18*, 339–356.

Shepsle, K. A., & Weingast, B. (1981). Structure-induced equilibrium and legislative choice. *Public Choice, 37*(3), 503–519.

Straffin, J., & Philip, D. (1980). *Topics in the theory of voting*. Boston: Birkhauser.

Postulates and Paradoxes of Voting Power in a Noncooperative Setting

Maria Montero

1 Introduction

Felsenthal and Machover (1998) define the P-power of a voter as the expected payoff the voter would get if the issue at stake is the division of a budget. They point out that the outcome of the bargaining process will not generally be deterministic,[1] and the index of P-power will be the average of the possible outcomes, weighted by their probability. This average could be computed by assigning a probability to each coalition S, and then a probability to each possible outcome conditional on S forming. They also point out that "A *bargaining model* would provide us with such probabilistic data, hence with a solution to the problem of measuring P-power. [...] The point is that no genuinely realistic general theoretical bargaining model is available" (Felsenthal and Machover, pp. 173 and 183).

While no theoretical bargaining model can be completely compelling, it may be argued that some of the existing models are more compelling than others. In particular, the equilibrium of the leading model of legislative bargaining, due to Baron and Ferejohn (1989), can be considered as a measure of P-power.[2] In this model, there is a budget to be divided and each voter has an equal probability of being selected to be the proposer. The proposer proposes a division of the budget

[1]They provide as an example the simple majority game with three players. If there was a unique deterministic outcome, symmetry points to the grand coalition with every player receiving $\frac{1}{3}$. However, this outcome seems too fragile. If we accept that a two-player coalition will eventually form, symmetry dictates that each of the three possible coalitions will be equally likely.

[2]Some of the papers on the Baron–Ferejohn model have referred to power indices as a benchmark for comparison (see Montero (2002, 2006), Snyder et al. (2005) and Kalandrakis (2006)).

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

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare,
DOI 10.1007/978-3-319-05158-1_17,

which is then voted upon. If the proposal passes, it is implemented; if not, a proposer is selected, again with equal probability for all players.

This paper discusses the equilibrium of this model from the point of view of some of the postulates and paradoxes in Chap. 7 of Felsenthal and Machover's book. It is well known that the equilibrium does not satisfy two of the minimal *adequacy postulates* that any reasonable measure of voting power must satisfy (ignoring dummies and vanishing just for dummies). Moreover, this paper shows that the equilibrium of the theoretical model does not satisfy another property that they consider essential: it does not respect *dominance*. It is possible for two voters to obtain the same payoff, even though one of the voters is strictly more desirable than the other. The equilibrium does satisfy a weaker version of dominance: a more desirable player cannot get a strictly lower payoff. Finally, the equilibrium of the theoretical model can display the *paradox of new members*. Whereas not all instances of the paradox of new members are truly paradoxical, it is argued that some of them are still surprising. The equilibrium is also compared to the Shapley (1953) value, which is the most widely accepted measure of P-power.

2 Preliminaries

2.1 Simple Games

Let $N = \{1, \dots, n\}$ be the set of players. $S \subseteq N$ ($S \neq \varnothing$) represents a generic coalition of players, and $v : 2^n \to \mathbb{R}$ with $v(\varnothing) = 0$ denotes the *characteristic function*. The (cooperative) game (N, v) is a *simple game* iff $v(S) \in \{0, 1\}$ for all $S \subseteq N$, $v(\varnothing) = 0$, $v(N) = 1$ and the following *monotonicity* condition is satisfied: $v(S) = 1$ implies $v(T) = 1$ for all S, T such that $S \subseteq T \subseteq N$. A coalition S is called *winning* iff $v(S) = 1$ and *losing* iff $v(S) = 0$. The set of winning coalitions is denoted by W. This set contains the same information as the function v. Indeed, Felsenthal and Machover (1998) denote the simple game as (N, W) rather than (N, v).

A coalition is *minimal winning* iff $v(S) = 1$ and $v(T) = 0$ for all T such that $T \subset S$. We will abbreviate minimal winning coalition as MWC.

A player such that $v(S \cup \{i\}) = v(S)$ for all S is called a *dummy player*. Dummy players do not belong to any MWC. A player who belongs to all winning coalitions is called a *veto player*.

A simple game is a *weighted majority game* iff there exist n nonnegative numbers (weights) $w_1, \dots, w_n$ and a positive number q such that $v(S) = 1$ if and only if $\sum_{i \in S} w_i := w(S) \geq q$. We will denote a weighted majority game by $[q; w_1, \dots, w_n]$.

2.2 The Baron–Ferejohn Model

Baron and Ferejohn's (1989) influential paper introduced a legislative bargaining game based on Rubinstein (1982) and Binmore (1987). In their paper n symmetric players must divide a budget by simple majority. Each player has an equal chance of being recognized to be the proposer; once a proposer is recognized he proposes a division of the budget. The rest of players then vote "yes" or "no"; if a majority of the players supports the proposal then it is implemented and the game ends; otherwise nature chooses a proposer again.

In extending the model to general voting games we must choose whether to keep the recognition probabilities identical for all players, or to have asymmetric probabilities. If the game is a weighted majority game, we may want to select each player with a probability proportional to his number of votes (this is done by Baron and Ferejohn in one of their examples). However, if we take this road, the equilibrium will not be exclusively a function of the set of winning coalitions W. For example, $[10; 9, 8, 1, 1]$ and $[3; 2, 1, 1, 1]$ both have the same set of winning coalitions but they have different expected payoffs if recognition probabilities are proportional to the weights.

Henceforth we will assume that each player has the same recognition probability.[3]

Formally, the bargaining model can be described in the following way. Let (N, v) be a simple game. Bargaining proceeds as follows. At every round $t = 1, 2, \ldots$ Nature selects a random proposer: player i is selected with probability $1/n$. This player proposes a distribution of the budget $(x_1, \ldots, x_n)$ with $x_j \geq 0$ for all $j = 1, \ldots n$ and $\sum_{j=1}^{n} x_j = 1$. The proposal is then voted upon. If the coalition of voters in favor of the proposal is winning, the proposal is implemented and the game ends; otherwise the game proceeds to the next period in which Nature selects a new proposer. Players are risk neutral and discount future payoffs by a factor $\delta \in [0, 1]$. A (pure) strategy for player i is a sequence $\sigma_i = (\sigma_i^t)_{t=1}^{\infty}$, where σ_i^t, the tth round strategy of player i, prescribes:

1. A proposal, denoted by x.
2. A response function assigning "yes" or "no" to all possible proposals by the other players.

Players may condition their actions on the history of play; however the literature focuses on equilibria in which they do not condition on any elements of history other than the current proposal, if any. The solution concept is *stationary subgame perfect equilibrium* (SSPE). Stationarity requires that players follow the same strategy at every round t regardless of past offers and responses to past offers. An SSPE always exists (Banks and Duggan 2000) and involves immediate agreement (Okada 1996).

[3] It is worth noting that, if the recognition probabilities coincide with the nucleolus (Schmeidler 1969), the nucleolus is the unique vector of expected payoffs (Montero 2006). Also, every payoff vector can be obtained for *some* vector of recognition probabilities (Kalandrakis 2006).

For $\delta < 1$, all SSPE lead to the same expected payoffs (Eraslan and McLennan 2013). We are interested in the case $\delta = 1$. We will take expected payoffs when $\delta = 1$ as a power measure; we will refer to this measure as the BF measure.[4]

The logic of the Baron–Ferejohn model of bargaining is very simple. Stationarity implies that a player's expected payoff given that a proposal is rejected does not depend on history; we will denote player i's expected payoff by y_i. Player i will be willing to accept any proposal that guarantees him at least y_i as a responder; as a proposer, he will convince the cheapest group of players whose votes are enough to form a winning coalition and will pay each of them exactly y_j. Formally, the proposer finds a coalition S that minimizes $\sum_{j \in S} y_j$ subject to the constraint that $S \cup \{i\}$ is a winning coalition.

Suppose there are three players, and decisions are taken by simple majority. Because strategies are stationary and players are symmetric, each player expects $1/3$ if a proposal is rejected. Each player is then prepared to accept any proposal that guarantees him at least $1/3$ as a responder. As a proposer, a player realizes he only needs to convince one other player and achieves this by offering the other player $1/3$ and keeping the remaining $2/3$ for himself. Agreement is immediate (even without discounting there is a pressure to reach an agreement in the first period because of the possibility of being excluded) and the proposer gets a disproportionate payoff. Also, coalitions are no greater than they need to be (no player is offered a positive payoff unless his vote is crucial for the proposal to be passed).

3 Postulates and Paradoxes

3.1 Felsenthal and Machover's Adequacy Postulates

A measure of voting power is a mapping ξ that assigns to any simple game W and any voter a of W a nonnegative real value $\xi_a[W]$. Felsenthal and Machover (henceforth FM) require the following three adequacy postulates:

(1) *Iso-invariance*: if there is an isomorphism of simple games from W to W' that maps a voter a to a', then $\xi_a[W] = \xi_{a'}[W']$.

The equilibrium of the bargaining game trivially satisfies this property.

(2) *Ignoring dummies*: If W and W' are two simple games that have exactly the same MWCs, then $\xi_a[W] = \xi_a[W']$ for any voter that is common to both.

Thus the addition of a dummy player cannot affect the payoff distribution between the other players. In particular, if the total payoff adds up to a constant,

[4]Equilibrium payoffs are usually unique even if $\delta = 1$. When they are not, one can take the limit of the expected equilibrium payoffs when $\delta \to 1$.

the dummy player must get 0. It is obvious that the equilibrium of the bargaining game does not satisfy this property: if we add a dummy to the three-player simple majority game, the dummy could obtain a positive payoff by offering two other players their continuation value. Since this is at most $\frac{2}{3}$, the dummy player can get at least $\frac{1}{4}\left(1-\frac{2}{3}\right) > 0$. The property would clearly hold if the bargaining model would assign a recognition probability of 0 to all dummy players.

(3) *Vanishing just for dummies*: $\xi_a[W] = 0$ iff a is a dummy in W.

This postulate is not satisfied in general because veto players must get everything (see Winter (1996) and Nohn (2013)). However, it is satisfied in the absence of veto players. This is because a player is guaranteed a nonnegative payoff as a responder and can always get a positive payoff as a proposer. Let i be the proposer, and j a voter with $y_j > 0$. Since j is not a veto player, i can always make a proposal in which all players other than j receive their continuation payoff (or slightly more) and i keeps the rest. Since $\sum_{k\neq j,i} y_k < 1$, this is positive.

The performance of the power index derived from the equilibrium of the theoretical model is very poor in terms of the adequacy postulates. However, all three postulates are satisfied (trivially in the case of postulate (2)) for games that have no veto or dummy players.

One may also question whether postulate (3) is that obvious. If there is a veto player and two nonveto players, one may argue that competition between the two nonveto players may drive their price to zero. Indeed, the only allocation in the core of the game would be $(1, 0, 0)$ so that postulate (3) is incompatible with core selection.

The BF model does not explicitly model coalition formation. However, the set of players who vote in favor of the final proposal can be considered as the coalition that forms. Then the BF measure is close to being based on MWCs only. If players have positive expected payoffs (which is always the case when there are no veto players), all coalition members other than the proposer must be pivotal, though the proposer may not be pivotal. Besides the trivial case of dummy players as proposers, Example 1 below illustrates that MWCs do not always form. Another exception to MWCs forming is games with veto players. Because all nonveto players get 0 in equilibrium, the coalition that forms may include more nonveto players than needed.

3.2 Dominance

We say that a dominates (or is more desirable than) b if $S \cup \{a\}$ is winning whenever $S \cup \{b\}$ is winning for all S such that $a \notin S$, $b \notin S$. The dominance relation is denoted by $a \succeq b$. We say that a strictly dominates (or is strictly more desirable than) b if $a \succeq b$ but not $b \succeq a$. This is denoted by $a \succ b$.

A power measure *respects dominance* if whenever $a \succ b$ in W then $\xi_a[W] > \xi_b[W]$ (Definition 7.6.6 in FM).

According to FM, any reasonable measure of P-power must respect dominance. They point out that the Shapley value satisfies the postulate but the Deegan and Packel (1978) index[5] violates it flagrantly (it is possible that $a \succ b$ but $\xi_a[W] < \xi_b[W]$).

The following proposition shows that the BF measure satisfies $y_i \geq y_j$ whenever $i \succeq j$, hence it never violates the postulate flagrantly. Example 1 then shows that it is possible for i to be strictly more desirable than j and nevertheless have the same payoff.[6]

Proposition 1. *Let (N, v) be a simple game and let i be more desirable than j. If i and j have the same probability of being proposer, then $y_i \geq y_j$.*

Proof. See appendix.■

Example 1. Consider the game with $N = \{1, 2, 3, 4, 5\}$ and minimal winning coalitions $\{1, 2\}, \{1, 3\}, \{2, 3, 4\}$ and $\{2, 3, 5\}$. This is the weighted game [7; 4, 3, 3, 1, 1]. Voter 1 is strictly more desirable than voters 2 and 3. However, the equilibrium payoff vector gives $\frac{3}{11}$ to the first three voters.

Proof. In order to show this, we construct strategies with two properties: first, the strategies must be optimal given that $\left(\frac{3}{11}, \frac{3}{11}, \frac{3}{11}, \frac{1}{11}, \frac{1}{11}\right)$ is the expected payoff vector; second, $\left(\frac{3}{11}, \frac{3}{11}, \frac{3}{11}, \frac{1}{11}, \frac{1}{11}\right)$ must be the expected payoff vector that results from playing the strategies.

Let $y = \left(\frac{3}{11}, \frac{3}{11}, \frac{3}{11}, \frac{1}{11}, \frac{1}{11}\right)$. The strategies we construct are as follows. As a responder, player i votes in favor of any proposal x with $x_i \geq y_i$. As a proposer, player i chooses the coalition S of minimal $\sum_{j \in S} y_j$ such that $S \cup \{i\}$ is winning. If $\left(\frac{3}{11}, \frac{3}{11}, \frac{3}{11}, \frac{1}{11}, \frac{1}{11}\right)$ is the expected payoff vector, player 1 has two optimal coalitions that he can propose: {1,2} and {1,3}. We assume without loss of generality that player 1 proposes each of the coalitions with equal probability. Players 4 and 5 are never of any use to player 1. As for player 2, the only optimal coalition is $\{1, 2\}$ ($\{2, 3, 4\}$ and $\{2, 3, 5\}$ are too expensive); similarly, $\{1, 3\}$ is the only optimal coalition for player 3. Player 4 is in only one MWC, $\{2, 3, 4\}$, but has another two winning coalitions that are equally cheap, $\{1, 2, 4\}$ and $\{1, 3, 4\}$. Let λ be the probability that player 4 proposes the MWC $\{2, 3, 4\}$; we assume that the other two coalitions are each proposed with probability $\frac{1-\lambda}{2}$. Similarly, player 5 proposes the MWC $\{2, 3, 5\}$ with probability λ and each of coalitions $\{1, 2, 5\}$ and $\{1, 3, 5\}$ with probability $\frac{1-\lambda}{2}$. If we can find a value of λ between 0 and 1 that induces y as the expected payoff vector, we have an equilibrium. It turns out that $\lambda = \frac{5}{6}$.

To see this, consider the expected payoffs for player 1 given the strategies. With probability $\frac{1}{5}$, player 1 is selected to be proposer. He then offers $\frac{3}{11}$ to either 2 or 3

[5] The Deegan-Packel index is calculated assuming that only MWCs form, each MWC has the same probability of forming, and members of the coalition that forms divide the payoff equally.

[6] Le Breton et al. (2012) also contains some games in which strict dominance is not respected, though the games are not proper (i.e. two disjoint coalitions can be winning).

and keeps $1 - \frac{3}{11}$. With probability $\frac{2}{5}$, either 2 or 3 is selected to be proposer and player 1 receives $\frac{3}{11}$. With probability $\frac{2}{5}$, either 4 or 5 is selected, and they offer $\frac{3}{11}$ to player 1 with probability $1 - \lambda = \frac{1}{6}$. Expected payoff for player 1 is then

$$\frac{1}{5}\left(1 - \frac{3}{11}\right) + \frac{2}{5}\frac{3}{11} + \frac{2}{5}\frac{1}{6}\frac{3}{11} = \frac{3}{11}.$$

As for player 2, he is selected to be proposer with probability $\frac{1}{5}$ and offers $\frac{3}{11}$ to player 1. With probability $\frac{1}{5}$, player 1 is selected to be proposer and, since player 1 randomizes between 2 and 3, player 2 receives $\frac{3}{11}$ with probability $\frac{1}{2}$. If player 3 is selected to be proposer player 2 receives nothing. With probability $\frac{1}{5}$, player 4 is selected to be proposer and proposes $\{2, 3, 4\}$ with probability $\frac{5}{6}$ and $\{1, 2, 4\}$ with probability $\frac{1}{12}$; in both cases player 2 receives $\frac{3}{11}$. The case in which 5 is selected is analogous: player 2 receives $\frac{3}{11}$ with probability $\left(\frac{5}{6} + \frac{1}{12}\right)$. Expected payoff for player 2 is then

$$\frac{1}{5}\left(1 - \frac{3}{11}\right) + \frac{1}{5}\frac{1}{2}\frac{3}{11} + \frac{2}{5}\left(\frac{5}{6} + \frac{1}{12}\right)\frac{3}{11} = \frac{3}{11}.$$

■

It is worth noting that the core does not respect dominance: voter 2 is strictly more desirable than voter 3 in the game $[7; 5, 2, 1, 1]$, but they both get 0 in the core.

3.3 The Paradox of New Members

The paradox of new members (Brams and Affuso 1976) occurs when enlargement of a voting body increases the power of an existing member even though the number of votes of all existing members and the quota remain constant. There are two ways in which the quota may remain constant: in absolute terms or in relative terms. Brams and Affuso consider both cases, whereas FM (footnote 3 in p. 235) insist in having the same quota in relative terms.

FM argue that the phenomenon is far from being paradoxical. Many of the instances of the paradox are not surprising, because they can be explained by the fact that the new member provides existing members with greater or easier possibilities of forming winning coalitions. In particular, it may be that a dummy player becomes a non-dummy player,[7] in which case the postulates of ignoring dummies and vanishing only for dummies require the paradox to occur. However, just because some instances of the paradox can be explained away it does not follow

[7]Kóczy (2009) shows that for all games with dummy players there is an enlargement such that a dummy player becomes nondummy.

that *all* instances of the paradox are reasonable. In particular, in one of the original examples of Brams and Affuso we see that a player gains from the addition of another player that does not seem of much use to him. This is the case in which $[4; 3, 2, 2]$ is enlarged to $[5; 3, 2, 2, 1]$.[8] It is clear that enlargement has hurt players 2 and 3 (before they could form a winning coalition on their own and now they need player 4). It is also clear that 1 has become the most powerful player in *relative* terms after enlargement, but it is not obvious why he should gain in *absolute* terms.[9]

Both the Shapley value and the Penrose measure display the paradox in this example. If we look at the Shapley value, we see that voter 1 becomes pivotal in a greater proportion of permutations of the players. Likewise, the Penrose measure of voter 1 increases because voter 1 is now pivotal for a greater proportion of vote configurations. However, the BF measure is based almost exclusively on minimal winning coalitions,[10] and voter 4 does not provide voter 1 with greater or easier possibilities of forming minimal winning coalitions. Perhaps surprisingly, the BF measure also displays the paradox in this example (see Drouvelis et al. 2010). Expected payoffs are $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$ in the first case and $\left(\frac{3}{8}, \frac{2}{8}, \frac{2}{8}, \frac{1}{8}\right)$ in the second case. Player 1 gains even though players 1 and 4 never include each other in the final coalition. This gain is partly due to players 2 and 3 always proposing to player 1 after enlargement. There is no obvious reason for this since coalition $\{2, 3, 4\}$ is equally cheap. The paradox has also been observed experimentally (Montero et al. 2008; Drouvelis et al. 2010) both under the BF protocol and under a more unstructured protocol. The emergence of the paradox in the experimental setting may be related to the lower transaction costs associated with smaller coalitions; note however that the theoretical bargaining model does not assume that smaller coalitions are inherently easier to form.

4 Comparison with the Shapley Value

The BF measure satisfies efficiency (Okada 1996, Theorem 1) and symmetry (Montero 2002, Lemma 2). As mentioned before, it does not satisfy the dummy player property. The fourth of Shapley's axioms, additivity, is not applicable for simple games because the sum of two simple games is not a simple game. Dubey

[8]In order to keep the quota constant in relative terms, $4/7 \times 8 \simeq 4.57$ votes would be required in the second game. A quota of 5 is equivalent to a quota of 4.57: any coalition that has at least 4.57 votes has at least 5 votes.

[9]The nucleolus payoff is unchanged.

[10]The only exceptions are cases in which the proposer is not pivotal, like Example 1, or cases in which many players are getting 0 because of the presence of a veto player, and since they are getting 0 they may as well be added to the coalition that forms.

(1975) replaced additivity by the transfer axiom. Laruelle and Valenciano (2001) show that this axiom is equivalent to another one that they call symmetric gain-loss. This property states that, if we compare a simple game v with the game v_S^* that results after deleting a minimal winning coalition $S \neq N$ from v, then the change in the Shapley value is the same for all players in S and for all players in $N \backslash S$. The following example illustrates this property:

Example 2. Consider the game $[5; 3, 2, 2, 1, 1]$. This game has the following minimal winning coalitions: $\{1,2\}, \{1,3\}, \{2,3,4\}, \{2,3,5\}, \{1,4,5\}$. If the game is modified so that coalition $\{1,4,5\}$ becomes losing, the three players in the coalition are equally affected according to the Shapley value but not according to the BF measure.

Note that the modified game is precisely the game analyzed in Example 1. The Shapley value of the original game is $(\frac{24}{60}, \frac{14}{60}, \frac{14}{60}, \frac{4}{60}, \frac{4}{60})$; after deleting coalition $\{1,4,5\}$ from the set of winning coalitions the Shapley value changes to $(\frac{22}{60}, \frac{17}{60}, \frac{17}{60}, \frac{2}{60}, \frac{2}{60})$. Thus, each of players 1, 4 and 5 have lost $\frac{2}{60}$. Equilibrium payoffs in the noncooperative model are approximately $(0.32, 0.22, 0.22, 0.12, 0, 12)$ for the first game and $\left(\frac{3}{11}, \frac{3}{11}, \frac{3}{11}, \frac{1}{11}, \frac{1}{11}\right)$ for the second game. It turns out that player 1 loses more than players 4 and 5: player 1's loss is approximately $0.32 - 0.27 = 0.05$, and player 4's loss is approximately $0.12 - 0.09 = 0.03$. There is no obvious reason why player 1 should lose more than the other two.[11]

Young (1985) characterized the Shapley value in the class of all games by iso-invariance, efficiency and the following property, called *marginality*:

We say that ξ satisfies marginality if, whenever v and w are two cooperative games with the same grand coalition and i is a player such that $v(S \cup \{i\}) - v(S) = w(S \cup \{i\}) - w(S)$ for every coalition S, then $\xi_i[v] = \xi_i[w]$.

Marginality means that a player's payoff only depends on his marginal contributions. The BF power index does not satisfy this property, as the following example illustrates:

Example 3. Consider the games [5;3,2,2,1] and [6;2,4,1,1]. Player 4 has the same marginal contributions to all coalitions in both games, and hence the same Shapley value, 1/12. Nevertheless, player 4's expected payoffs are 1/8 and 0 respectively (the value of 0 arises because player 2 is a veto player in the second game).

It is easy to see that player 4 has the same marginal contributions in both games. All singleton coalitions are losing in both games, and remain losing after adding player 4. Coalition $\{1,2\}$ is winning in both games. Coalition $\{1,3\}$ is winning in the first game and losing in the second game (but remains losing after adding player 4). Coalition $\{2,3\}$ is losing and becomes winning after adding player 4 in both games. Finally, coalition $\{1,2,3\}$ is already winning in both games.

[11] Indeed, the nucleolus moves from $\left(\frac{3}{9}, \frac{2}{9}, \frac{2}{9}, \frac{1}{9}, \frac{1}{9}\right)$ to $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0\right)$; see Montero (2005).

The BF measure for the game $[5; 3, 2, 2, 1]$ is $\left(\frac{3}{8}, \frac{2}{8}, \frac{2}{8}, \frac{1}{8}\right)$. Given these payoffs, player 1 has two optimal coalitions, $\{1, 2\}$ and $\{1, 3\}$ (player 4 is of no use to player 1); each of the two coalitions is proposed with probability $\frac{1}{2}$. Players 2 and 3 are indifferent between proposing a coalition with player 1 and proposing $\{2, 3, 4\}$; in equilibrium they always propose a coalition with player 1. Player 4 has only one optimal coalition. Given that player 4 receives no proposals, its payoff is given by $\frac{1}{4}\left[1 - \frac{4}{8}\right] = \frac{1}{8}$.

In the second game, player 1 proposes $\{1, 2\}$ and players 3 and 4 propose $\{2, 3, 4\}$. Player 2 will propose $\{1, 2\}$ or $\{2, 3, 4\}$ depending on whether $\{1\}$ or $\{3, 4\}$ is cheaper. Then $y_2 = \frac{1}{4}[1 - y(S)] + \frac{3}{4}y_2$, where S is either $\{1\}$ or $\{3, 4\}$. We see that y_2 is a weighted average of itself and $1 - y(S)$, thus $y_2 = 1 - y(S)$, or $y_2 + y(S) = 1$. Hence the excluded players have an expected payoff of 0. But since player 2 is including the cheapest players in the coalition, players in S must also have $y(S) = 0$, hence $y_2 = 1$ and $y_4 = 0$.

5 Concluding Remarks

The BF measure does badly in terms of FM's postulates of voting power. It may be argued that some of those failures are not too serious. Non-dummy players can only get 0 in the presence of veto players, and we may be willing to sacrifice the postulate "vanishing only for dummies" in favor of core selection in this case. The positive payoff of dummy players is more problematic, though it could be easily (if arbitrarily) eliminated by modifying the measure in such a way that dummy players cannot make proposals. This paper has shown that the measure does not respect dominance, and this failure cannot be justified by core selection (the core is empty in the example provided). Finally, a player may gain from enlargement according to this measure even if it never forms a coalition with the new member.

Appendix: Proof of Proposition 1

Let θ_i denote the probability that i is selected to be the proposer. We will show that $i \succeq j$ and $\theta_i = \theta_j$ imply $y_i \geq y_j$ for an arbitrary θ_i and δ; the BF measure is a particular case in which $\theta_i = \theta_j = \frac{1}{n}$ and $\delta \to 1$.

By contradiction, suppose $y_i < y_j$ in equilibrium. Equilibrium strategies may not be unique, though equilibrium payoffs are (Eraslan and McLennan 2013). Fix a combination of equilibrium strategies. Let S_i^* be any of the coalitions that are optimal for player i to propose in equilibrium. Let λ_{ij} be the probability that i proposes to j (i.e., offers j its continuation value) in this particular equilibrium. Expected equilibrium payoffs for player i satisfy the equation

$$y_i = \theta_i \left[1 - \delta \sum_{k \in S_i^* \setminus \{i\}} y_k \right] + \sum_{k \in N \setminus \{i\}} \theta_k \lambda_{ki} \delta y_i .$$

Expected equilibrium payoffs for j are defined analogously.

In order to compare y_i and y_j, it will be helpful to re-arrange this equation. If we add and subtract $\theta_i \delta y_i$ from the right-hand side and collect terms, we can write $y_i = \theta_i \left[1 - \delta \sum_{k \in S_i^*} y_k\right] + \sum_{k \in N \setminus \{i\}} \theta_k \lambda_{ki} \delta y_i + \theta_i \delta y_i$. We can write $\sum_{k \in N \setminus \{i\}} \theta_k \lambda_{ki}$ as r_i (this is the probability that i receives a proposal from another player) and $1 - \delta \sum_{k \in S_i^*} y_k$ as π_i (this is the payoff i gets as a proposer over and above its continuation value).

This yields $y_i = \theta_i \pi_i + r_i \delta y_i + \theta_i \delta y_i$. Solving for y_i we find

$$y_i = \frac{\theta_i \pi_i}{1 - r_i \delta - \theta_i \delta} \tag{A.1}$$

We now compare this expression for i and j and reach a contradiction.

In the numerator we have $\theta_i \pi_i$. Clearly, $\theta_i = \theta_j$ by assumption. As for $\pi_i := 1 - \delta \sum_{k \in S_i^*} y_k$, each proposer proposes one of the cheapest winning coalitions to which they belong. Because i can replace j in any winning coalition of which j is a member, the optimal coalition for i is at least as cheap as the optimal coalition for j, thus $\pi_i \geq \pi_j$.

In the denominator, we need to compare r_i and r_j. If we look at players k other than i and j, we know that $\lambda_{ki} \geq \lambda_{kj}$ for all $k \neq i, j$. This is because third parties will either include both i and j in their proposed coalition, only i, or none of them (if a player is including j but not i, j could be replaced by i, and the coalition would still be winning and cheaper). As for i and j themselves, suppose j is not proposing to i for sure. This would mean that j belongs to a winning coalition T such that $i \notin T$ and $\sum_{k \in T} y_k \leq \sum_{k \in S} y_k$ for all S such that $S \in W$ and $i, j \in S$. Because $T \setminus \{j\} \cup \{i\}$ is available to i and $y_i < y_j$, i would never propose to j. Thus, either j is proposing to i for sure (in which case clearly $\theta_j \lambda_{ji} \geq \theta_i \lambda_{ij}$) or i never proposes to j (which also implies $\theta_j \lambda_{ji} \geq \theta_i \lambda_{ij}$). Either way, $\theta_j \lambda_{ji} \geq \theta_i \lambda_{ij}$ and overall $r_i \geq r_j$.

Since the numerator is at least as large and the denominator is at least as small for y_i compared to y_j, it follows that $y_i \geq y_j$, contradicting our initial assumption $y_i < y_j$.

The only case not covered by this proof occurs if Eq. (A.1) is not valid because the denominator would be 0. This can only happen if $\delta = 1$ and $\theta_i + r_i = 1$, implying that i always belongs to the coalition that forms. Expected equilibrium payoffs in this case are given by

$$y_i = \theta_i \left[1 - \sum_{k \in S_i^*} y_k \right] + y_i .$$

It follows that $1 - \sum_{k \in S_i^*} y_k = 0$, i.e. the coalition S_i^* involves all the players who receive a positive payoff in equilibrium. This is impossible if there are no veto players. In the absence of veto players, any individual player $k \neq i$ with $y_k > 0$ could be dropped from the coalition so no coalition S_i^* with $\sum_{k \in S_i^*} y_k = 1$ could be optimal for i unless $y_i = 1$, but then i would be dropped from j's proposed coalition, contradicting $y_i = 1$. If there are veto players, there are multiple equilibria for $\delta = 1$ but in the limit when $\delta \to 1$ we have $y_i = \theta_i$, hence $y_i = y_j$ (Nohn 2013).

References

Banks, J. S., & Duggan, J. (2000). A bargaining model of collective choice. *American Political Science Review*, *94*, 73–88.

Baron, D. P., & Ferejohn, J. A. (1989). Bargaining in legislatures. *American Political Science Review*, *83*, 1181–1206.

Binmore, K. (1987). Perfect equilibria in bargaining models. In K. Binmore, & P. Dasgupta (Eds.), *The economics of bargaining* (pp. 77–105). Oxford: Blackwell.

Brams, S. J., & Affuso, P. J. (1976). Power and size: a new paradox. *Theory and Decision*, *7*, 29–56.

Deegan, J., & Packel, E. W. (1978). A new index of power for simple n-person games. *International Journal of Game Theory*, *7*, 113–123.

Drouvelis, M., Montero, M., & Sefton, M. (2010) Gaining power through enlargement: strategic foundations and experimental evidence. *Games and Economic Behavior, 69*(2), 274–292.

Dubey, P. (1975). On the uniqueness of the Shapley value. *International Journal of Game Theory*, *4*, 131–140.

Eraslan, H., & McLennan, A. (2013) Uniqueness of stationary equilibrium payoffs in coalitional bargaining. *Journal of Economic Theory, 148*, 2195–2222.

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

Kalandrakis, T. (2006) Proposal rights and political power. *American Journal of Political Science*, *50*, 441–448.

Kóczy, L. (2009). Measuring voting power: the paradox of new member vs. the null player axiom. In I. J. Rudas et al. (Eds.), *Towards intelligent engineering and information technology*. SCI 243 (pp. 67–78). Heidelberg: Springer.

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

Le Breton, M., Montero, M., & Zaporozhets, V. (2012). Voting power in the EU Council of Ministers and fair decision making in distributive politics. *Mathematical Social Sciences, 63*, 159–173.

Montero, M. (2002). Noncooperative bargaining in apex games and the kernel. *Games and Economic Behavior*, *41*, 309–321.

Montero, M. (2005). On the nucleolus as a power index. *Homo Oeconomicus*, *4*, 551–567.

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

Montero, M., Sefton, M., & Zhang, P. (2008). Enlargement and the balance of power: an experimental study. *Social Choice and Welfare*, *30*, 69–87.

Nohn, A. (2013). Veto players and noncooperative foundations of power in coalitional bargaining. In M.J. Holler and H. Nurmi (Eds.) Power, Voting and Voting Power: 30 Years After, Springer-Verlag Berlin Heidelberg.

Okada, A. (1996) A noncooperative coalitional bargaining game with random proposers. *Games and Economic Behavior*, *16*, 97–108.
Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica*, *50*, 97–108.
Schmeidler, D. (1969). The nucleolus of a characteristic function game. *SIAM Journal on Applied Mathematics*, *17*, 1163–1170.
Shapley, L.S. (1953). A value for n-person games. *Annals of Mathematical Studies*, *28*, 307–317.
Snyder, J. M., Ting, M. M., & Ansolabehere, S. (2005). Legislative bargaining under weighted voting. *American Economic Review*, *95*, 981–1004.
Winter, E. (1996). Voting and vetoing. *American Political Science Review*, *90*, 813–823.
Young, H. P. (1985). Monotonic solutions of cooperative games. *International Journal of Game Theory*, *14*, 65–72.

Satisfaction Approval Voting

Steven J. Brams and D. Marc Kilgour

1 Introduction

Approval voting (AV) is a voting system in which voters can vote for, or approve of, as many candidates as they like. Each approved candidate receives one vote, and the candidates with the most votes win.

This system is well suited to electing a single winner, which almost all the literature on AV since the 1970s has addressed (Brams and Fishburn 1978, 1983/2007; Brams 2008, chs. 1 and 2). But for multiwinner elections, such as for seats on a council or in a legislature, AV's selection of the most popular candidates or parties can fail to reflect the diversity of interests in the electorate.

As a possible solution to this problem when voters use an *approval ballot*,[1] in which they can approve or not approve of each candidate, we propose *satisfaction approval voting* (SAV). SAV works as follows when the candidates are individuals. A voter's *satisfaction score* is the fraction of his or her approved candidates who are elected, whether the voter is relatively discriminating (i.e., approves of few candidates) or not (approves of many candidates). In particular, it offers a strategic

[1]Merrill III and Nagel (1987) were the first to distinguish between *approval balloting*, in which voters can approve of one or more candidates, and approval voting (AV), a method for aggregating approval ballots. SAV, as we will argue, is a method of aggregation that tends to elect candidates in multiwinner elections who are more representative of the entire electorate than those elected by AV, who are simply the most popular candidates.

S.J. Brams (✉)
Department of Politics, New York University, 19 W. 4th Street, New York, NY 10012, USA
e-mail: steven.brams@nyu.edu

D.M. Kilgour
Department of Mathematics, Wilfrid Laurier University, Waterloo, ON, Canada N2L 3C5
e-mail: mkilgour@wlu.ca

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare, DOI 10.1007/978-3-319-05158-1_18,

choice to voters, who may bullet vote (i.e., exclusively for one candidate) or vote for several candidates, perhaps hoping to make a specific set of candidates victorious.

Among all the sets of candidates that might be elected, SAV chooses the set that maximizes the *sum* of all voters' satisfaction scores. As we will show, SAV may give very different outcomes from AV; SAV outcomes are not only more satisfying to voters but also tend to be more representative of the diversity of interests in an electorate.[2] Moreover, they are easy to calculate.

In Sect. 2, we apply SAV to the election of individual candidates (e.g., to a council) when there are no political parties. We show, in the extreme, that SAV and AV may elect disjoint subsets of candidates. When they differ, SAV winners will generally represent the electorate better—by at least partially satisfying more voters—than AV winners. While maximizing total voter satisfaction, however, SAV may not maximize the *number* of voters who approve of at least one winner—one measure of representativeness—though it is more likely to do so than AV.

This is shown empirically in Sect. 3, where SAV is applied to the 2003 Game Theory Society (GTS) election of 12 new Council members from a list of 24 candidates (there were 161 voters). SAV would have elected two winners different from the 12 elected under AV and would have made the Council more representative of the entire electorate. We emphasize, however, that GTS members might well have voted differently under SAV than under AV, so one cannot simply extrapolate a reconstructed outcome, using a different aggregation method, to predict the consequences of SAV.

In Sect. 4, we consider the conditions under which, in a 3-candidate election with 2 candidates to be elected, a voter's ballot might change the outcome, either by making or breaking a tie. In our decision-theoretic analysis of the 19 contingencies in which this is possible, approving of one's two best candidates induces a preferred outcome in about the same number of contingencies as bullet voting, even though a voter must split his or her vote when voting for 2 candidates. More general results on optimal voting strategies under SAV are also discussed.

In Sect. 5, we apply SAV to party-list systems, whereby voters can approve of as many parties as they like. Parties nominate their "quotas," which are based on their vote shares, rounded up; they are allocated seats to maximize total voter satisfaction, measured by the fractions of nominees from voters' approved parties that are elected. We show that maximizing total voter satisfaction leads to the proportional representation (PR) of parties, based on the Jefferson/d'Hondt method of apportionment, which favors large parties.

SAV tends to encourage multiple parties to share support, because they can win more seats by doing so. At the same time, supporters of a party diminish its

[2]Representing this diversity is not the issue when electing a single winner, such as a mayor, governor, or president. In such an election, the goal is to find a consensus choice, and we believe that AV is better suited than SAV to satisfy this goal. Scoring rules, in which voters rank candidates and scores are associated with the ranks, may also serve this end, but the optimal scoring rule for achieving particular standards of justice (utilitarianism, maximin, or maximax) is sensitive to the distribution of voter utilities (Apesteguia et al. 2011).

individual support by approving of other parties, so there is a trade-off between helping a favorite party and helping a coalition of parties that may be able to win more seats *in toto*. Some voters may want to support only a favorite party, whereas others may want to support multiple parties that, they hope, will form a governing coalition. We argue that this freedom is likely to make parties more responsive to the wishes of their supporters with respect to (1) other parties with which they coalesce and (2) the candidates they choose to nominate.[3]

In Sect. 6, we conclude that SAV may well induce parties to form coalitions, if not merge, *before* an election. This will afford voters the ability better to predict what policies the coalition will promote, if it forms the next government, and, therefore, to vote more knowledgeably.[4] In turn, it gives parties a strong incentive to take careful account of their supporters' preferences, including their preferences for coalitions with other parties.

2 Satisfaction Approval Voting for Individual Candidates

We begin by applying SAV to the election of individual candidates, such as to a council or legislature, in which there are no political parties. We assume in the subsequent analysis that there are at least two candidates to be elected, and more than this number run for office (to make the election competitive).

To define SAV formally, assume that there are $m > 2$ candidates, numbered 1, 2, $\ldots$, m. The set of all candidates is $\{1, 2, \ldots, m\} = [m]$, and k candidates are to be elected, where $2 \leq k < m$. Assume voter i approves of a subset of candidates $V_i \subseteq [m]$, where $V_i \neq \emptyset$. (Thus, a voter may approve of only 1 candidate, though more are to be elected.) For any subset of k candidates, S, voter i's satisfaction is $\frac{|V_i \cap S|}{|V_i|}$, or the fraction of his or her approved candidates that are elected.[5] SAV elects a subset of k candidates that maximizes

[3] The latter kind of responsiveness would be reinforced if voters, in addition to being able to approve of one or more parties, could use SAV to choose a party's nominees.

[4] More speculatively, SAV may reduce a multiparty system to two competing coalitions of parties. The majority coalition winner would then depend, possibly, on a centrist party that can swing the balance in favor of one coalition or the other. Alternatively, a third moderate party (e.g., Kadima in Israel) might emerge that peels away supporters from the left and the right. In general, SAV is likely to make coalitions more fluid and responsive to popular sentiment.

[5] An interesting modification of this measure was suggested by Kilgour and Marshall (2011) to apply when a voter approves of more candidates than are to be elected: Change the denominator of the satisfaction measure from $|V_i|$ to $\min\{|V_i|, k\}$. Thus, for example, if voter i approves of 3 candidates, but only $k = 2$ can be elected, i's satisfaction would be 2/2 (rather than 2/3) whenever any two of his or her approved candidates are elected. This modification ensures that a voter's influence on the election is not diluted if he or she approves of more candidates than are to be elected, but it does not preserve other properties of SAV.

$$s(S) = \sum_i \frac{|V_i \cap S|}{|V_i|}, \tag{1}$$

which we interpret as the *total satisfaction* of voters for S. By convention, $s(\varnothing) = 0$.

To illustrate SAV, assume there are $m = 4$ candidates, $\{a, b, c, d\}$, and 10 voters who approve of the following subsets[6]:

4 voters: *ab*
3 voters: *c*
3 voters: *d*.

Assume $k = 2$ of the 4 candidates are to be elected. AV elects $\{a, b\}$, because a and b receive 4 votes each, compared to 3 votes each that c and d receive. By contrast, SAV elects $\{c, d\}$, because the satisfaction scores of the six different two-winner subsets are as follows:

$$s(a,b) = 4(1) = 4$$
$$s(a,c) = s(a,d) = s(b,c) = s(b,d) = 4\left(\frac{1}{2}\right) + 3(1) = 5$$
$$s(c,d) = 3(1) + 3(1) = 6$$

Thus, the election of c and d gives 6 voters full satisfaction of 1, which corresponds to greater total satisfaction, 6, than achieved by the election of any other pair of candidates.[7]

A *candidate's* satisfaction score—as opposed to a voter's satisfaction score—is the sum of the satisfaction scores of voters who approve of him or her. For example, if a candidate receives 3 votes from bullet voters, 2 votes from voters who approve of two candidates, and 5 votes from voters who approve of three candidates, his or her satisfaction score is 3(1) + 2(½) + 5(1/3) = 5 2/3.

More formally, candidate j's *satisfaction score* is $s(j) = \sum_i \frac{|V_i \cap j|}{|V_i|}$, whereas candidate j's *approval score* is $a(j) = \sum |V_i \cap j|$. Our first proposition shows that satisfaction scores make it easy to identify all winning subsets of candidates under SAV—that is, all subsets that maximize total satisfaction.

Proposition 1. *Under SAV, the k winners are any k candidates whose individual satisfaction scores are the highest.*

---

[6] We use *ab* to indicate the *strategy* of approving of the subset $\{a, b\}$, but we use $\{a, b\}$ to indicate the *outcome* of a voting procedure. Later we drop the set-theoretic notation, but the distinction between voter strategies and election outcomes is useful for now.

[7] Arguably, candidates c and d benefit under SAV by getting bullet votes from their supporters. While their supporters do not share their approval with other candidates, their election gives representation to a majority of voters, whereas AV does not.

Proof. Because $V_i \cap S = \bigcup_{j \in S} (V_i \cap j)$, it follows from (1) that

$$s(S) = \sum_i \left(\frac{1}{|V_i|}\right) \sum_{j \in S} |V_i \cap j| = \sum_{j \in S} \sum_i \frac{|V_i \cap j|}{|V_i|} = \sum_{j \in S} s(j).$$

Thus, the satisfaction score of any subset S, $s(S)$, can be obtained by summing the satisfaction scores of the individual members of S. Now suppose that $s(j)$ has been calculated for all candidates $j = 1, 2, \ldots, m$. Then, for any arrangement of the set of candidates $[m]$ so that the scores $s(j)$ are in non-increasing order, the first k candidates constitute a subset of candidates that maximizes total voter satisfaction.□

As an illustration of Proposition 1, consider the previous example, in which

$$s(a) = s(b) = 4\left(\frac{1}{2}\right) = 2$$
$$s(c) = s(d) = 3(1) = 3$$

Because c and d have higher satisfaction scores than any other candidates, the subset $\{c, d\}$ is the unique winning subset if $k = 2$ candidates are to be elected under SAV.

One consequence of Proposition 1 is a characterization of *tied elections*: There are two or more winning subsets if and only if the satisfaction scores of the k^{th} and $(k+1)^{\text{st}}$ candidates are tied in satisfaction score when the candidates are arranged in descending order, as described in the proof of Proposition 1. This follows from the fact that tied subsets must contain the k most satisfying candidates, but if those in the k^{th} and the $(k+1)^{\text{st}}$ positions give the same satisfaction, a subset containing either would maximize total voter satisfaction. Ties among three or more sets of candidates are, of course, also possible.

It is worth noting that the satisfaction that a voter gains when an approved candidate is elected does not depend on how many of the voter's other approved candidates are elected, as some multiple-winner systems that use an approval ballot prescribe.[8] This renders candidates' satisfaction scores *additive*: The satisfaction from electing subsets of two or more candidates is the sum of the candidates'

[8]Two of these systems—proportional AV and sequential proportional AV—assume that a voter's satisfaction is marginally decreasing—the more of his or her approved candidates are elected, the less satisfaction the voter derives from having additional approved candidates elected. See http://www.nationmaster.com/encyclopedia/Proportional-approval-voting; http://www.nationmaster.com/encyclopedia/Sequential-proportional-approval-voting for a description and examples of these two systems, and Alcalde-Unzu and Vorsatz (2009) for an axiomatic treatment of systems in which the points given to a candidate are decreasing in the number of candidates of whom the voter approves, which they call "size approval voting." More generally, see Kilgour (2010) and Kilgour and Marshall (2011) for a comparison of several different approval-ballot voting systems that have been proposed for the election of multiple winners, all of which may give different outcomes.

satisfaction scores. Additivity greatly facilitates the determination of SAV outcomes when there are multiple winners—simply choose the subset of individual candidates with the highest satisfaction scores.

The additivity of candidate satisfaction scores reflects SAV's equal treatment of voters: Each voter has one vote, which is divided evenly among all his or her approved candidates. Thus, if two candidates are vying for membership in the elected subset, then gaining the support of an additional voter always increases a candidate's score by $1/x$, where x is the number of candidates approved of by that voter.[9] This is a *consequence* of the goal of maximizing total voter satisfaction, not an assumption about how approval votes are to be divided.

We next compare the different outcomes that AV and SAV can induce.

Proposition 2. *AV and SAV can elect disjoint subsets of candidates.*

Proof. This is demonstrated by the previous example: AV elects $\{a, b\}$, whereas SAV elects $\{c, d\}$.

For any subset S of the candidates, we say that S *represents* a voter i if and only if voter i approves of some candidate in S. We now ask how representative is the set of candidates who win under SAV or AV—that is, how many voters approve of at least one elected candidate.

SAV winners usually represent at least as many, and often more, voters than the set of AV winners, as illustrated by the previous example, in which SAV represents 6 voters and AV only 4 voters. SAV winners c and d appeal to distinctive voters, who are more numerous and so win under SAV, whereas AV winners a and b appeal to the same voters but, together, receive more approval and so win under AV.

But there are (perhaps unlikely) exceptions:

Proposition 3. *An AV outcome can be more representative than a SAV outcome.*

Proof. Assume there are $m = 5$ candidates and 13 voters, who vote as follows:

2 voters: *a*
5 voters: *ab*
6 voters: *cde.*

If 2 candidates are to be elected, the AV outcome is either $\{a, c\}$, $\{a, d\}$, or $\{a, e\}$ (7 approvals for a, and 6 each for c, d, and e), whereas the SAV outcome is $\{a, b\}$, because

[9]By contrast, under *cumulative voting* (*CV*), a voter can divide his or her votes—or, equivalently, a single vote—unequally, giving more weight to some candidates than others. However, *equal and even cumulative voting* (*EaECV*), which restricts voters to casting the same number of votes for all candidates whom they support, is equivalent to SAV, though its connection to voter satisfaction, as far as we know, has not previously been demonstrated. While CV and EaECV have been successfully used in some small cities in the United States to give representation to minorities on city councils, it seems less practicable in large elections, including those in countries with party-list systems in which voters vote for political parties (Sect. 5). See http://en.wikipedia.org/wiki/Cumulative_voting for additional information on cumulative voting.

$$s(a) = 2(1) + 5\left(\frac{1}{2}\right) = 4\frac{1}{2}$$
$$s(b) = 5\left(\frac{1}{2}\right) = 2\frac{1}{2}$$
$$s(c) = s(d) = s(e) = 6\,(1/3) = 2.$$

Thus, whichever of the three AV outcomes is selected, the winning subset represents all 13 voters, whereas the winners under SAV represent only 7 voters.□

The "problem" for SAV in the forgoing example would disappear if candidates c, d, and e were to combine forces and became one candidate (say, c), rendering $s(c) = 6(1) = 6$. Then the SAV and AV outcomes would both be $\{a, c\}$, which would give representation to all 13 voters. Indeed, as we will show when we apply SAV to party-list systems in Sect. 5, SAV encourages parties to coalesce to increase their combined seat share.

But first we consider another possible problem of both SAV and AV.

Proposition 4. *There can be subsets that represent more voters than either the SAV or the AV outcome.*

Proof. Assume there are $m = 5$ candidates and 12 voters, who vote as follows:

4 voters: *ab*
4 voters: *acd*
3 voters: *ade*
1 voter: *e*.

If 2 candidates are to be elected, the AV outcome is $\{a, d\}$ (11 and 7 votes, respectively, for a and d), and the SAV outcome is also $\{a, d\}$, because

$$s(a) = 4\left(\frac{1}{2}\right) + 7\,(1/3) = 4\ 1/3$$
$$s(b) = 4\left(\frac{1}{2}\right) = 2$$
$$s(c) = 4\,(1/3) = 1\ 1/3$$
$$s(d) = 7\,(1/3) = 2\ 1/3$$
$$s(e) = 3\,(1/3) + 1(1) = 2.$$

While subset $\{a, d\}$ represents 11 of the 12 voters, subset $\{a, e\}$ represents all 12 voters.□

Interestingly enough, the so-called *greedy algorithm* (for representativeness) would select $\{a, e\}$. It works as follows. The candidate who represents the most voters—the AV winner—is selected first. Then the candidate who represents as many of the remaining (unrepresented) voters as possible is selected next, then the candidate who represents as many as possible of the voters not represented by the first two candidates is selected, and so on. The algorithm ends as soon as all voters are represented, or until the required number of candidates is selected. In the example used to prove Proposition 4, the greedy algorithm first chooses candidate a (11 votes) and then candidate e (1 vote).

Given a set of ballots, we say a *minimal representative set* is a subset of candidates with the properties that (1) every voter approves at least one candidate in the subset, and (2) there are no smaller subsets with property (1). In general, finding a minimal representative set is computationally difficult.[10] Although the greedy algorithm finds a minimal representative set in the previous example, it is no panacea.

Proposition 5. *SAV can find a minimal representative set when both AV and the greedy algorithm fail to do so.*

Proof. Assume there are $m = 3$ candidates and 17 voters, who vote as follows:

5 voters: ab
5 voters: ac
4 voters: b
3 voters: c.

If 2 candidates are to be elected, the AV outcome is $\{a, b\}$ (a gets 10 and b gets 9 votes), which is identical to the subset produced by the greedy algorithm.[11] On the other hand, the SAV outcome is $\{b, c\}$, because

$$s(a) = 5\left(\frac{1}{2}\right) + 5\left(\frac{1}{2}\right) = 5$$

$$s(b) = 5\left(\frac{1}{2}\right) + 4(1) = 6\frac{1}{2}$$

$$s(c) = 5\left(\frac{1}{2}\right) + 3(1) = 5\frac{1}{2}.$$

Not only does this outcome represent all 17 voters, but it is also the minimal representative set.□

The greedy algorithm fails to find the minimal representative set in the previous example because it elects the "wrong" candidate—the AV winner, a—first.

[10]Technically, the problem is NP hard (http://en.wikipedia.org/wiki/NP-hard), because it is equivalent to the hitting-set problem, which is a version of the vertex-covering problem (http://en.wikipedia.org/wiki/Vertex_cover) discussed in Karp (1972). Under SAV, as we showed at the beginning of this section, the satisfaction-maximizing subset of, say, k candidates can be calculated efficiently, as it must contain only candidates with satisfaction scores among the k highest. Because of this feature, the procedure is practical for multiwinner elections with many candidates.

[11]Candidates a, b, and c receive, respectively, 10, 9, and 8 votes; the greedy algorithm first selects a (10 votes) and then b (4 votes).

Curiously, a closely related example shows that none of these methods may find a minimal representative subset:

Proposition 6. *SAV, AV, and the greedy algorithm can all fail to find a unique minimal representative set.*

Proof. Assume there are $m = 3$ candidates and 9 voters, who vote as follows:

3 voters: *ab*
3 voters: *ac*
2 voters: *b*
1 voters: *c*

AV and the greedy algorithm give $\{a, b\}$, as in the previous example, but so does SAV because

$$s(a) = 3\left(\frac{1}{2}\right) + 3\left(\frac{1}{2}\right) = 3$$
$$s(b) = 3\left(\frac{1}{2}\right) + 2(1) = 3\frac{1}{2}$$
$$s(c) = 3\left(\frac{1}{2}\right) + 1(1) = 2\frac{1}{2}.$$

As before, $\{b, c\}$ is the minimal representative set.□

Minimal representative sets help us assess and compare SAV and AV outcomes of elections; the greedy algorithm contributes by finding an upper bound on the size of a minimal representative set, because it eventually finds a set that represents all voters, even if it is not minimal. But there is a practical problem with basing an election procedure on the minimal representative set: Only by chance will that set have k members. If it is either smaller or larger, it must be "adjusted."

But what adjustment is appropriate? For example, if the minimal representative set is too small, should one add candidates that give as many voters as possible a second representative, then a third, and so on? Or, after each voter has approved of at least one winner, should it, like SAV, maximize total voter satisfaction? It seems to us that maximizing total voter satisfaction from the start is a simple and desirable goal, even if it sometimes sacrifices some representativeness.

Another issue, addressed in the next proposition, is vulnerability to candidate cloning. AV is almost defenseless against cloning, whereas SAV exhibits some resistance.[12]

A *clone* of a candidate is a new candidate who is approved by exactly the supporters of the original candidate. We call a candidate, h, a *minimal winning* candidate (under AV or SAV) if the score of every other winning candidate is at

[12] AV-related systems, like proportional AV and sequential proportional AV (see note 8), seem to share AV's vulnerability, but we do not pursue this question here.

least equal to the score of h; otherwise, h is a *nonminimal winning* candidate. We consider whether a clone of a winning candidate is certain to be elected; if so, a minimal winning candidate will be displaced.

We say that a winning candidate can clone successfully if its clone is certain to be elected. For any two candidates j and h, denote the set of voters who support both j and h by $V(j,h) = \{i : j \in V_i, h \in V_i\}$, and denote the set of voters who support j but not h by $V(j,-h) = \{i : j \in V_i, h \notin V_i\}$.

Proposition 7. *Under AV, any nonminimal winning candidate can clone successfully. Under SAV, a nonminimal winning candidate, j, cannot clone successfully if and only if, for every winning candidate* $h \neq j$,

$$\sum_{i \in V(j,-h)} \frac{1}{|V_i| + 1} < \sum_{i \in V(h,-j)} \frac{1}{|V_i|} < \sum_{i \in V(j,-h)} \frac{1}{|V_i|}$$

Proof. Suppose that j, a non-minimal winning candidate under AV, clones. After cloning, the approval scores of all original candidates, including j, are unchanged, and the approval score of j's clone is the same as j's. Therefore, both j and its clone have approval scores that exceed that of the original minimal candidate(s), and both j and the clone will belong to winning set, to the exclusion of an original minimal winning candidate.

Now suppose that j, a nonminimal winning candidate under SAV, clones. Clearly, j succeeds at cloning if and only if j and its clone displace some winning candidate h whose satisfaction score is necessarily less than $s(j)$. For such a candidate, h, we must have

$$s(j) = \sum_{i \in V(j,h)} \frac{1}{|V_i|} + \sum_{i \in V(j,-h)} \frac{1}{|V_i|} > s(h) = \sum_{i \in V(h,j)} \frac{1}{|V_i|} + \sum_{i \in V(h,-j)} \frac{1}{|V_i|}$$

or, in other words,

$$\sum_{i \in V(h,-j)} \frac{1}{|V_i|} < \sum_{i \in V(j,-h)} \frac{1}{|V_i|}.$$

Let $s^n(j)$ and $s^n(h)$ be the satisfaction scores of j and h after cloning. If cloning fails to displace h, it must be the case that

$$\begin{aligned} s^n(h) &= \sum_{i \in V(h,j)} \frac{1}{|V_i| + 1} + \sum_{i \in V(h,-j)} \frac{1}{|V_i|} \\ &> \sum_{i \in V(j,h)} \frac{1}{|V_i| + 1} + \sum_{i \in V(j,-h)} \frac{1}{|V_i| + 1} = s^n(j) \end{aligned}$$

or, in other words,

$$\sum_{i \in V(j,-h)} \frac{1}{|V_i| + 1} < \sum_{i \in V(h,-j)} \frac{1}{|V_i|},$$

which is easily seen to complete the proof.□

Note that the second inequality of Proposition 7 is equivalent to $s(j) > s(h)$, which means that the original satisfaction score of h must be less than the original satisfaction score of j, so that the clone displaces a lower-ranked candidate.

To see that the condition of Proposition 7 has bite, consider an example with $m = 4$ candidates and 17 voters—who are to elect 2 candidates—and vote as follows:

6 voters: *ab*
6 voters: *ac*
5 voters: *d*

Under SAV, the scores are $s(a) = 6$, $s(b) = s(c) = 3$, and $s(d) = 5$, so the winning subset is $\{a, d\}$. If a clones, then both a and its clone have satisfaction scores of 4, whereas the score of d remains 5, so d is not displaced by a's clone, and cloning is unsuccessful.

We conclude that, relative to AV, SAV discourages the formation of clones unless a candidate's support is sufficiently large that he or she can afford to transfer a substantial part of it to a clone and still win—in which case the clone, as well as the original candidate, would both seem deserving of election.

We turn next to a real election, in which AV was used to elect multiple winners, and assess the possible effects of SAV, had it been used. We are well aware that voters might have voted differently under SAV and take up this question in Sect. 4.

3 The Game Theory Society Election

In 2003, the Game Theory Society (GTS) used AV for the first time to elect 12 new council members from a list of 24 candidates. (The council comprises 36 members, with 12 elected each year to serve 3-year terms.[13]) We give below the numbers of members who voted for from 1 to all 24 candidates (no voters voted for between 19 and 23 candidates):

[13]The fact that there is exit from the council after 3 years makes the voting incentives different from a society in which (1) members, once elected, do not leave and (2) members decide who is admitted (Barberà et al. 2001).

Votes cast	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	24
Number of voters	3	2	3	10	8	6	13	12	21	14	9	25	10	7	6	5	3	3	1

Casting a total of 1,574 votes, the 161 voters, who constituted 45 % of the GTS membership, approved, on average, $1,574/161 \approx 9.8$ candidates; the median number of candidates approved of, 10, is almost the same.[14]

The modal number of candidates approved of is 12 (by 25 voters), echoing the ballot instructions that 12 of the 24 candidates were to be elected. The approval of candidates ranged from a high of 110 votes (68.3 % approval) to a low of 31 votes (19.3 % approval). The average approval received by a candidate was 40.7 %. Because the election was conducted under AV, the elected candidates were the 12 most approved, who turned out to be all those who received at least 69 votes (42.9 % approval). Do these AV winners best represent the electorate? With the caveat that the voters might well have approved of different candidates if SAV rather than AV had been used, we compare next how the outcome would have been different if SAV had been used to aggregate approval votes.

Under SAV, 2 of the 12 AV winners would not have been elected.[15] Each set of winners is given below—ordered from most popular on the left to the least popular on the right, as measured by approval votes—with differences between those who were elected under AV and those who would have been elected under SAV underscored:

$$\text{AV}: 111111111\underline{1}1\underline{100}0000000000$$
$$\text{SAV}: 111111111\underline{0}1\underline{011}0000000000$$

Observe that the AV winners who came in 10th (70 votes) and 12th (69 votes) would have been displaced under SAV by the candidates who came in 13th (66 votes) and 14th (62 votes), according to AV, and just missed out on being elected.

Recall that a voter is *represented* by a subset of candidates if he or she approves of at least one candidate in that subset. The elected subset under SAV represents all but 2 of the 161 voters, whereas the elected subset under AV failed to represent 5 of the 161 voters. But neither of these subsets is the best possible; the greedy algorithm gives a subset of 9 candidates that represents all 161 voters, which includes 5 of the

[14]Under SAV, whose results we present next, the satisfaction scores of voters in the GTS election are almost uncorrelated with the numbers of candidates they approved of, so the number of candidates approved of does not affect, in general, a voter's satisfaction score—at least if he or she had voted the same as under AV (a big "if" that we investigate later).

[15]Under the "minimax procedure" (Brams et al. 2007; Brams 2008), 4 of the 12 AV winners would not have been elected. These 4 include the 2 who would not have been elected under SAV; they would have been replaced by 2 who would have been elected under SAV. Thus, SAV partly duplicates the minimax outcome. It is remarkable that these two very different systems agree, to an extent, on which candidates to replace to make the outcome more representative.

AV winners and 6 SAV winners, including the 2 who would have won under SAV but not under AV.

It turns out, however, that this is not a minimal representative set of winners: There are more than a dozen subsets of 8 candidates, though none of 7 or fewer candidates, that represent all 161 voters, making 8 the minimal size of a representative set.[16] To reduce the number of such sets, it seemed reasonable to ask which one maximizes the minimum satisfaction of all 161 voters.

This criterion, however, was not discriminating enough to produce one subset that most helped the least-satisfied voter: There were 4 such subsets that gave the least-satisfied voter a satisfaction score of $1/8 = 0.125$—that is, that elected one of his or her approved candidates. To select the "best" among these, we used as a second criterion the one that maximizes total voter satisfaction, which gives

$$100111000000110001000001.$$

Observe that only 4 of the 8 most approved candidates are selected; moreover, the remaining four candidates include the least-approved candidate (24th on the list).

But ensuring that every voter approves of at least one winner comes at a cost. The total satisfaction that the aforementioned minimal representative set gives is 60.9, whereas the subset of 8 candidates that maximizes total voter satisfaction—without regard to giving every voter an approved representative—is

$$111110011000100000000000.$$

Observe that six of the eight most approved candidates are selected (the lowest candidate is 13th on the list). The total satisfaction of this subset is 74.3, which is a 22 % increase over the above score of the most satisfying minimal representative set. We leave open the question whether such an increase in satisfaction is worth the disenfranchisement of a few voters.

In choosing a minimal representative set, the size of an elected voting body is allowed to be endogenous. In fact, it could be as small as one candidate if one candidate is approved of by everybody.

By contrast, if the size of the winning set is fixed, then once a minimal representative set has been selected—if that is possible—then one can compute the larger-than-minimal representative set that maximizes total voter satisfaction. In the case of the GTS, because there is a minimal representative set with only 8 members, we know that a 12-member representative set is certainly feasible.

In making SAV and related calculations for the GTS election, we extrapolated from the AV ballots. We caution that our extrapolations depend on the assumption that GTS voters would not have voted differently under SAV than under AV. In particular, under SAV, would GTS voters have been willing to divide their one

[16]We are grateful to Richard F. Potthoff for writing an integer program that gave the results for the GTS election that we report on next.

vote among multiple candidates if they thought that their favorite candidate needed their undivided vote to win?

4 Voting for Multiple Candidates Under SAV: A Decision-Theoretic Analysis

To try to answer the foregoing question, we begin by analyzing a much simpler situation—there are 3 candidates, with 2 to be elected. As shown in Table 1, there are exactly 19 contingencies in which a single voter's strategy can be *decisive*—that is, make a difference in which 2 of the 3 candidates are elected—by making or breaking a tie among the candidates. In decision theory, these contingencies are the so-called states of nature.

In Table 1, the contingencies are shown as the numbers of votes that separate the three candidates.[17] For example, contingency 4 (1, ½, 0) indicates that candidate a is ahead of candidate b by ½ vote, and that candidate b is ahead of candidate c by ½ vote.[18] The outcomes produced by a voter's strategies in the left column of Table 1 are indicated either (1) by the two candidates elected (e.g., ab), (2) by a candidate followed by two candidates who tie for second place, indicated by a slash (e.g., $a - b/c$), or (3) by all the candidates in a three-way tie ($a/b/c$).

A voter may choose any one of the six strategies by approving of either one or two candidates. (Approving of all three candidates, or none at all, would have no effect on the outcome, so we exclude them as strategies that can be decisive.[19]) To determine the optimal strategies of a voter, whom we call the *focal voter*, we posit that he or she has strict preference $a \succ b \succ c$.

We assume that the focal voter has preferences not only for individual candidates but also over sets of two or three candidates. In particular, given this voter's strict

[17]Notice that the numbers of votes shown in a contingency are all within 1 of each other, enabling a voter's strategy to be decisive; these numbers need not sum to an integer, even though the total number of voters and votes sum to an integer. For example, contingency 4 can arise if there are 2 ab voters and 1 ac voter, giving satisfaction scores of 3/2, 1, and ½, respectively, to a, b, and c, which sum to 3. But this is equivalent to contingency 4 (1, ½, 0), obtained by subtracting ½ from each candidate's score, whose values do not sum to an integer. Contingencies of the form (1, ½, ½), while feasible, are not included, because they are equivalent to contingencies of the form (½, 0, 0)—candidate a is ½ vote ahead of candidates b and c.

[18]We have not shown contingencies in which any candidate is guaranteed a win or a loss. The 19 contingencies in Table 1 represent all states in which the strategy of a voter can make each of the three candidates a winner or a loser, rendering them *3-candidate competitive contingencies*.

[19]If there were a minimum number of votes (e.g., a simple majority) that a candidate needs in order to win, then abstention or approving of everybody could matter. But here we assume the two candidates with the most votes win, unless there is a tie, in which case we assume there is an (unspecified) tie-breaking rule.

Table 1 Strategies and outcomes for 19 contingencies in 3-candidate, 2-winner elections in which one voter can be decisive

Contingency										
Voter	1	2	3	4	5	6	7	8	9	10
Strategy	1, 1, 0	1, 0, 1	0, 1, 1	1, ½, 0	1, 0, ½	½, 1, 0	0, 1, ½	½, 0, 1	0, ½, 1	1, 0, 0
a	ab^*	ac^*	$\underline{a-b/c}^*$	ab^*	ac	ab^*	$\underline{ab}^*$	ac^*	ac^*	$a-b/c$
ab	ab^*	ac^*	bc	ab^*	$a-b/c$	ab^*	$b-a/c$	ac^*	bc	ab^*
b	ab^*	$a/b/c^*$	bc	ab^*	$\underline{ab}^*$	ab^*	bc	bc	bc	ab^*
bc	ab^*	ac^*	bc	ab^*	ac	$b-a/c$	bc	$c-a/b$	bc	$a-b/c$
c	$a/b/c$	ac^*	bc	ac	ac	bc	bc	ac	bc	ac
ac	ab^*	ac^*	bc	$a-b/c$	ac	ab^*	bc	ac	$c-a/b$	ac
Contingency										
Voter	11	12	13	14	15	16	17	18	19	
Strategy	0, 1, 0	0, 0, 1	½, ½, 0	½, 0, ½	0, ½, ½	½, 0, 0	0, ½, 0	0, 0, ½	0, 0, 0	
a	ab^*	ac^*	ab^*	ac	$a-b/c^*$	$a-b/c$	ab^*	ac^*	$a-b/c$	
ab	ab^*	$c-a/b$	ab^*	$\underline{a-b/c}^*$	$b-a/c$	ab^*	ab^*	$a/b/c^*$	$\underline{ab}^*$	
b	$b-a/c$	bc	ab^*	$b-a/c$	bc	ab^*	$b-a/c$	bc	$b-a/c$	
bc	bc	bc	$b-a/c$	$c-a/b$	bc	$a/b/c$	bc	bc	bc	
c	bc	$c-a/b$	$c-a/b$	ac	bc	ac	bc	$c-a/b$	$c-a/b$	
ac	$b-a/c$	ac^*	$a-b/c$	ac	$c-a/b$	ac	$a/b/c$	ac^*	ac	

Note: The outcomes produced by a voter's strategies in the left columns of this table are indicated (1) by the two candidates elected (e.g., ab), (2) by a candidate followed by two candidates who tie for second place, separated by a slash (e.g., $a-b/c$), or (3) by the candidates in a three-way tie ($a/b/c$). For the focal voter with preference $a \succ b \succ c$, starred outcomes indicate his or her best or tied-for-best outcomes for each contingency; underscores indicate a uniquely best outcome.

preference for individual candidates, we assume the following preference relations for pairs and triples of candidates:

$$ab \succ \; a - b/c \succ ac \approx b - a/c \approx a/b/c \succ c - a/b \succ bc,$$

where "$\approx$" indicates indifference, or a tie, between pairs of outcomes: Neither outcome in the pair is strictly better than the other. Thus, the certain election of *a* and *c* (*ac*) is no better nor worse than either the certain election of *b* and the possible election of either *a* or *c* ($b - a/c$), or the possible election of any pair of *a*, *b*, or *c* ($a/b/c$).[20]

We have starred the outcomes, for each contingency, that are the best or the tied-for-best for the focal voter; underscores indicate a *uniquely* best outcome. In contingency 4, for example, there are four starred *ab* outcomes, all of which give the focal voter's top two candidates. These outcomes are associated with the focal voter's first four strategies; by contrast, his or her other two strategies elect less preferred sets of candidates.

In contingency 7, outcome *ab*, associated with the focal voter's strategy *a*, is not only starred but also underscored, because it is a uniquely best outcome. A strategy that is associated with a uniquely best outcome is *weakly undominated*, because no other strategy can give at least as good an outcome for that contingency.

Observe from Table 1 that strategy *a* leads to a uniquely best outcome in 4 contingencies (3, 7, 9, and 15), strategy *ab* in 2 contingencies (14 and 19), and strategy *b* in 1 contingency (5), rendering all these strategies weakly undominated. It is not difficult to show that the focal voter's other three strategies, all of which involve approving of *c*, are *weakly dominated*:

- *a*, *ab*, and *b* weakly dominate *bc*
- *a* and *ab* weakly dominate *c*
- *a* weakly dominates *ac*.

In no contingency does a weakly dominated strategy lead to a better outcome than a strategy that dominates it, and in at least one contingency it leads to a strictly worse outcome.

Among the weakly undominated strategies, *a* leads to at least a tied-for-best outcome in 14 contingencies, *ab* in 13 contingencies (9 of the *a* and *ab* contingencies overlap), and *b* in 8 contingencies. In sum, it is pretty much a toss-up between weakly undominated strategies *a* and *ab*, with *b* a distant third-place finisher.

[20]Depending on the tie-breaking rule, the focal voter may have strict preferences over these outcomes, too. Because each allows for the possibility of any pair of winning candidates, we chose not to distinguish them. To be sure, $a - b/c$ (second best) and $c - a/b$ (second worst) also allow for the possibility of any pair of winning candidates, but the fact that the first involves the certain election of *a*, and the second the certain election of *c*, endows them with, respectively, a more-preferred and less-preferred status than the three outcomes among which the focal voter is indifferent.

It is no fluke that the focal voter's three strategies that include voting for candidate *c* (*c*, *ac*, and *bc*) are all weakly dominated.

Proposition 8. *If there is more than one candidate, a strategy that includes approving of a least-preferred candidate is weakly dominated under SAV.*

Proof. Let W be a focal voter's strategy that includes approving of a least-preferred ("worst") candidate, w. Let $\overline{W}$ be the focal voter's strategy of duplicating W, except for approving of w, unless W involves voting only for w. In that case, let $\overline{W}$ be a strategy of voting for any candidate other than w.

Assume that the focal voter chooses $\overline{W}$. Then $\overline{W}$ will elect the same candidates that W does except, possibly, for w. However, there will be at least one contingency in which $\overline{W}$ does not elect w with certainty (e.g., in a contingency in which w is assigned 0) and W does, but none in which the reverse is the case. Hence, $\overline{W}$ weakly dominates W.□

In Table 1, voting for a second choice, candidate *b*, is a weakly undominated strategy, because it leads to a uniquely best outcome in contingency 5. This is not the case for AV, in which a weakly undominated strategy includes always approving of a most-preferred candidate—not just never approving of a least-preferred candidate (Brams and Fishburn 1978).

Thus, SAV admits more weakly undominated strategies than AV. In some situations, it may be in the interest of a voter to approve of set of strictly less-preferred candidates and forsake a set of strictly more-preferred candidates. As a case in point, assume a focal voter strictly ranks 5 candidates as follows, $a \succ b \succ c \succ d \succ e$, and 2 candidates are to be elected. In contingency $(a, b, c, d, e) = (0, 0, 3/4, 1, 1)$, strategy *ab* elects candidates *d* and *e*, the focal voter's two worst choices, whereas strategy *cd*, comprising less-preferred candidates, elects candidates *c* and *d*, which is a strictly better outcome.

To conclude, our decision-theoretic analysis of the 3-candidate, 2-winner case demonstrates that voting for one's two most-preferred candidates leads to the same number of uniquely best and about the same number of at least tied-for-best outcomes, despite the fact that voters who vote for more than one candidate must split their votes evenly under SAV. We plan to investigate whether this finding carries over to elections in which there are more candidates and more winners, as well as the effect that the ratio of candidates to winners has.

Unlike AV, approving of just a second choice when there are 3 competitive candidates is a weakly undominated strategy under SAV, though it is uniquely optimal in only one of the 19 contingencies.[21] More generally, while it is never optimal for a focal voter to select a strategy that includes approving of a worst candidate

[21] To the degree that voters have relatively complete information on the standing of candidates (e.g., from polls), they can identify the most plausible contingencies and better formulate optimal strategies, taking into account the likely optimal strategies of voters with opposed preferences. In this situation, a game-theoretic model would be more appropriate than a decision-theoretic model for analyzing the consequences of different voting procedures. We plan to investigate such models

(not surprising), sometimes it is better to approve of strictly inferior candidates than strictly superior candidates (more surprising), though this seems relatively rare.

5 Voting for Political Parties

In most party-list systems, voters vote for political parties, which win seats in a parliament in proportion to the number of votes they receive. We now propose a SAV-based party voting system in which voters would not be restricted to voting for one party but could vote for as many parties as they like. If a voter approves of x parties, each approved party's score would increase by $1/x$.

Unlike standard apportionment methods, some of which we will describe shortly, our SAV system does *not* award seats according to the *quota* to which a party is entitled. (A party's quota is a number of seats such that its proportion of the seats is exactly equal to the proportion of its supporters in the electorate. Note that a quota is typically not an integer.) Instead, parties are allocated seats to maximize total voter satisfaction, measured by the fractions of nominees from voters' approved parties that are elected.

We begin our discussion with an example, after which we formalize the application of SAV to party-list systems. Then we return to the example to illustrate the possible effects of voting for more than one party.

5.1 Bullet Voting

Effectively, SAV requires that the number of candidates nominated by a party equal its *upper quota* (its quota rounded up). To illustrate, consider the following 3-party, 11-voter example, in which three seats are to be filled (we indicate parties by capital letters).

5 voters support A
4 voters support B
2 voters support C.

Assume that the supporters of each party vote exclusively for it. Party I's *quota*, q_i, is its proportion of votes times 3, the number of seats to be apportioned:

$$q_A = (5/11)(3) \approx 1.364$$
$$q_B = (4/11)(3) \approx 1.091$$
$$q_C = (2/11)(3) \approx 0.545.$$

in the future. For models of strategic behavior in proportional-representation systems—but not those that use an approval ballot—see Slinko and White (2010).

Under SAV, each party is treated as if it had nominated a number of candidates equal to its upper quota, so *A*, *B*, and *C* have effectively nominated 2, 2, and 1 candidates, respectively—2 more than the number of candidates to be elected. We emphasize that the numbers of candidates nominated are not a choice that the parties make but follow from their quotas, based on the election returns.

SAV finds apportionments of seats to parties that (1) maximize total voter satisfaction and (2) are *monotonic*: A party that receives more votes than another cannot receive fewer seats.

In our previous example, there are three monotonic apportionments to parties (*A*, *B*, *C*)—(3, 0, 0), (2, 1, 0) and (1, 1, 1)—giving satisfaction scores of

$$s\,(3,0,0) = 5(1) + 4(0) + 2(0) = 5$$
$$s\,(2,1,0) = 5(1) + 4\left(\frac{1}{2}\right) + 2(0) = 7$$
$$s\,(1,1,1) = 5\left(\frac{1}{2}\right) + 4\left(\frac{1}{2}\right) + 2(1) = 6\frac{1}{2}.$$

Apportionment (2, 1, 0) maximizes the satisfaction score, giving

- 5 *A* voters satisfaction of 1 for getting *A*'s 2 nominees elected
- 4 *B* voters satisfaction of ½ for getting 1 of *B*'s 2 nominees elected
- 2 C voters satisfaction of 0, because C's nominee is not elected.

5.2 Formalization

In a SAV election of k candidates from lists provided by parties $1, 2, \ldots, p$, suppose that party j has v_j supporters, and that $\sum_{j=1}^{p} v_j = n$. Then party j's *quota* is $q_j = \frac{v_j}{n}k$. If q_j is an integer, party j is allocated exactly q_j seats.

We henceforth assume that all parties' quotas are nonintegral. Then party j receives either its *lower quota*, $l_j = \lfloor q_j \rfloor$, or its *upper quota*, $u_j = \lceil q_j \rceil$. Of course, $u_j = l_j + 1$. In total, $r = k - \sum_{j=1}^{p} l_j$ parties receive their upper quota rather than their lower quota. By assumption, $r > 0$. The set of parties receiving upper quota, $S \subseteq [p] = \{1, 2, \ldots, p\}$, is chosen to maximize the total satisfaction of all voters, $s(S)$, subject to $|S| = r$.

Recall that when electing individual candidates, SAV chooses *candidates* that maximize total voter satisfaction. When allocating seats to parties, SAV finds *apportionments* of seats that maximize total voter satisfaction.

The apportionment in our example is *not* an apportionment according to the *Hamilton method* (also called "largest remainders"), which begins by giving each party the integer portion of its exact quota (1 seat to A and 1 seat to B). Then any remaining seats go to the parties with the largest remainders until the seats are exhausted, which means that C, with the largest remainder (0.545), gets the third seat, yielding the apportionment (1, 1, 1) to (A, B, C).

There are five so-called divisor methods of apportionment (Balinski et al. 1982/2001). Among these, only the Jefferson/d'Hondt method, which favors larger parties, gives the SAV apportionment of (2, 1, 0) in our example.[22] This is no accident, as shown by the next proposition.

Proposition 9. *The SAV voting system for political parties gives the same apportionment as the Jefferson/d'Hondt apportionment method, but with an upper-quota restriction.*[23] *SAV apportionments also satisfy lower quota and thus satisfy quota.*

Proof. Each of party j's v_j voters gets satisfaction of 1 if party j is allocated its upper quota, and satisfaction $\frac{l_j}{u_j}$ if party j is allocated its lower quota. If the subset of parties receiving upper quota is $S \subseteq [p]$, then the total satisfaction over all voters is

$$s(S) = \sum_{j \in S} v_j + \sum_{j \notin S} v_j \left(\frac{l_j}{u_j}\right) = \sum_{j=1}^{p} v_j - \sum_{j \notin S} \frac{v_j}{u_j}, \tag{2}$$

where the latter equality holds because $\frac{l_j}{u_j} = 1 - \frac{1}{u_j}$. The SAV apportionment is, therefore, determined by choosing S such that $|S| = r$ and S maximizes $s(S)$, which by (2) can be achieved by choosing $S^c = [p] - S$ to minimize $\sum_{j \in S^c} \frac{v_j}{u_j}$. Clearly, this requirement is achieved when S contains the r largest values of $\frac{v_j}{u_j}$.

To compare the SAV apportionment with the Jefferson/d'Hondt apportionment, assume that all parties have already received l_j seats. The first party to receive u_j seats is, according to Jefferson/d'Hondt, the party, j, that maximizes $\frac{v_j}{l_j+1} = \frac{v_j}{u_j}$. After this party's allocation has been adjusted to equal its upper quota, remove it from the set of parties. The next party to receive u_j according to Jefferson/d'Hondt is the remaining party with the greatest value of $\frac{v_j}{u_j}$, and so on. Clearly, parties

[22]The *Jefferson/d'Hondt method* allocates seats sequentially, giving the next seat to the party that maximizes $v/(a+1)$, where v is its number of voters and a is its *present* apportionment. Thus, the 1st seat goes to A, because $5 > 4 > 2$ when $a = 0$. Now $a = 1$ for A and remains 0 for B and C. Because $4/1 > 5/2 > 2/1$, B gets the second seat. Now $a = 1$ for A and B and remains 0 for C. Because $5/2 > 4/2 = 2/1$, A gets the third seat, giving an apportionment of (2, 1, 0) to (A, B, C). The divisor method that next-most favors large parties is the *Webster/Sainte-Laguë* method, under which the party that maximizes $v/(a + ½)$ gets the next seat. After A and B get the first two seats, the third seat goes to C, because $2/(½) > 5/(3/2) > 4/(3/2)$, so the Webster/Sainte-Laguë method gives an apportionment of (1, 1, 1) to (A, B, C).

[23]There are objective functions with a min/max operator that Jefferson/d'Hondt also optimizes (Balinski and Young 1982/2001, p. 105; Fernández de Córdoba and Penandés 2009), but they are more difficult to justify in the context of seat assignments.

receive seats in decreasing order of their values of $\frac{v_j}{u_j}$. Because Jefferson/d'Hondt apportionments always satisfy lower quota (Balinski et al. 1982/2001, pp. 91 and 130), SAV apportionments satisfy quota (i.e., both upper and lower).[24]□

A consequence of this procedure is that SAV apportionments are certain to satisfy upper quota, unlike (unrestricted) Jefferson/d'Hondt apportionments. Effectively, parties cannot nominate candidates for, and therefore cannot receive, more seats than their quotas rounded up.[25]

Because SAV produces Jefferson/d'Hondt apportionments, except for the upper-quota restriction, SAV favors large parties. Nevertheless, small parties will not be wiped out, provided their quotas are at least 1, assuming that no threshold, or minimum vote to qualify for a seat, is imposed (in some countries, the threshold is 5 % or more of the total vote).

5.3 Multiple-Party Voting

If a voter votes for multiple parties, his or her vote is equally divided among all his or her approved parties. To illustrate in our previous example, suppose parties B and C reach an agreement on policy issues, so that their 6 $(4+2)$ supporters approve of both parties. Meanwhile, the 5 party A supporters continue to vote for A alone.

Now the vote totals of B and C are taken to equal $6(½) = 3$, making the quotas of the three parties the following:

$$\begin{aligned} q_A &= (5/11)\,(3) \approx 1.364 \\ q_B &= (3/11)\,(3) \approx 0.818 \\ q_C &= (3/11)\,(3) \approx 0.818. \end{aligned}$$

By the algorithm above, party seats are allocated in decreasing order of $\frac{v_j}{u_j}$. Because these ratios are $5/2 = 2.5$, $3/1 = 3.0$, and $3/1 = 3.0$ for parties A, B, and

---

[24]The Jefferson/d'Hondt method with an upper-quota constraint is what Balinski and Young (1982/2001, p. 139) call Jefferson-Quota; SAV effectively provides this constraint. Balinski and Young (1982/2001, ch. 12) argue that because it is desirable that large parties be favored and coalitions encouraged in a parliament, the Jefferson/d'Hondt method should be used, but they do not impose the upper-quota constraint that is automatic under SAV. However, in earlier work (Balinski and Young 1978), they—along with Still (1979)—looked more favorably on such a constraint.

[25]This SAV-based system could be designed for either a closed-list or an open-list system of proportional representation. In a closed-list system, parties would propose an ordering of candidates prior to the election; the results of the election would tell them how far down the list they can go in nominating their upper quotas of candidates. By contrast, in an open-list system, voters could vote for individual candidates; the candidates' vote totals would then determine their positions on their party lists.

C respectively, it follows that the apportionment of seats is (1, 1, 1). Compared with apportionment (2, 1, 0) earlier with bullet voting, *A* loses a seat, *B* stays the same, and *C* gains a seat.

In general, parties that are too small to be represented at all cannot hurt themselves by approving of each other. However, the strategy may either help or hurt the combined seat count of parties that achieve at least one seat on their own. In the previous example, *B* and *C* supporters together ensure themselves of a majority of 2 seats if they approve of each other's party, but they may nonetheless choose to go their separate ways.

One reason is that *B* does not *individually* benefit from supporting *C*; presumably, B's supporters would need to receive some *collective* benefit from supporting *C* to make it worth their while also to approve of C. Note that if only 2 of B's supporters also approve C, but both of C's supporters approve of B, the vote counts $(5, 2 + 4/2, 4/2) = (5, 4, 2)$, would be exactly as they were originally, so the outcome of the election would be unchanged.

A possible way around this problem is for *B* and *C* to become one party, assuming that they are ideologically compatible, reducing the party system to just two parties. Because the combination of *B* and *C* has more supporters than *A* does, this combined party would win a majority of seats.

6 Conclusions

We have proposed a new voting system, satisfaction approval voting (SAV), for multiwinner elections. It uses an approval ballot, whereby voters can approve of as many candidates or parties as they like, but they do not win seats based on the number of approval votes they receive.

We first considered the use of SAV in elections in which there are no political parties, such as in electing members of a city council. SAV elects the set of candidates that maximizes the satisfaction of all voters, where a voter's satisfaction is the fraction of his or her approved candidates who are elected. This measure works equally well for voters who approve of few or of many candidates and, in this sense, can mirror a voter's personal tastes.

A candidate's satisfaction score is the sum of the satisfactions that his or her election would give to all voters. Thus, a voter who approves of a candidate contributes $1/x$ to the candidate's satisfaction score, where x is the total number of candidates of whom the voter approves. The winning set of candidates is the one with the highest individual satisfaction scores.

Among other findings, we showed that SAV and AV may elect disjoint sets of candidates. SAV tends to elect candidates that give more voters either partial or complete satisfaction—and thus representation—than does AV, but this is not universally true and is a question that deserves further investigation.

Additionally, SAV inhibits candidates from creating clones to increase their representation. But voting for a single candidate can be seen as risky for a voter, as the voter's satisfaction score will be either 0 or 1, so risk-averse voters may be inclined to approve of multiple candidates.

SAV may not elect a representative set of candidates—whereby every voter approves of at least one elected candidate—as we showed would have been the case in the 2003 election of the Game Theory Society Council. However, the SAV outcome would have been more representative than the AV outcome (given the approval ballots remained the same as in the AV election). Yet we also showed that a fully representative outcome could have been achieved with a smaller subset of candidates (8 instead of 12).

Because SAV divides a voter's vote evenly among the candidates he or she approves of, SAV may encourage more bullet voting than AV does. However, we found evidence that, in 3-candidate, 2-winner competitive elections, voters would find it almost equally attractive to approve of their two best choices as their single best choice. Unlike AV, they may vote for strictly less-preferred candidates if they think their more-preferred candidates cannot benefit from their help.

We think the most compelling application of SAV is to party-list systems. Each party would provide either an ordering of candidates, or let the vote totals for individual candidates determine this ordering. Each party would then be considered to have *nominated* a number of candidates equal to its upper quota after the election. The set of candidates *elected* would be any one that maximizes total voter satisfaction among monotonic apportionments.

Because parties nominate, in general, more candidates than there are seats to be filled, not every voter can be completely satisfied. We showed that the apportionment of seats to parties under SAV gives the Jefferson/d'Hondt apportionment method with a quota constraint, which tends to favor larger parties while still ensuring that all parties receive at least their lower quotas.

To analyze the effects of voting for multiple parties, we compared a scenario in which voters bullet voted with a scenario in which they voted for multiple parties. Individually, parties are hurt when their supporters approve of other parties. Collectively, however, they may be able to increase their combined seat share by forming coalitions—whose supporters approve all parties in it—or even by merging. At a minimum, SAV may discourage parties from splitting up unless to do so would mean they would be able to recombine to form a new and larger party, as Kadima did in Israel.

Normatively speaking, we believe that better coordination by parties should be encouraged, because it would give voters a clearer idea of what to expect when they decide which parties to support—compared to the typical situation today, when voters can never be sure what parties will join in a governing coalition and what its policies will be. Because this coordination makes it easier for voters to know what parties to approve of, and for party coalitions to form that reflect their supporters' interests, we believe that SAV is likely to lead to more informed voting and more responsive government in parliamentary systems.

Acknowledgments We thank Joseph N. Ornstein and Erica Marshall for valuable research assistance, and Richard F. Potthoff for help with computer calculations that we specifically acknowledge in the article. We are also grateful for the helpful comments of an anonymous reviewer.

References

Alcalde-Unzu, J., & Vorsatz, M. (2009). Size approval voting. *Journal of Economic Theory, 144*(3), 1187–1210.

Apesteguia, J., Ballester, M. A., & Ferrer, R. (2011). On the justice of decision rules. *Review of Economic Studies, 78*(1), 1–16.

Balinski, M. L., & Young, H. P. (1978). Stability, coalitions and schisms in proportional representation systems. *American Political Science Review, 72*(3), 848–858.

Balinski, M. L., & Young, H. P. (1982/2001). *Fair representation: Meeting the ideal of one man, one vote*. New Haven/Washington, DC: Yale University Press/Brookings Institution.

Barberà, S., Maschler, M., & Shalev, J. (2001). Voting for voters: A model of electoral evolution. *Games and Economic Behavior, 37*(1), 40–78.

Brams, S. J. (2008). *Mathematics and democracy: Designing better voting and fair-division procedures*. Princeton, NJ: Princeton University Press.

Brams, S. J., & Fishburn, P. C. (1978). Approval voting. *American Political Science Review, 72*(3), 831–847.

Brams, S. J., & Fishburn, P. C. (1983/2007). *Approval voting*. New York: Springer.

Brams, S. J., Kilgour, D. M., & Sanver, M.R. (2007). A minimax procedure for electing committees. *Public Choice, 132*(3–4), 401–420.

Fernández de Córdoba, G., & Penandés, A. (2009). Institutionalizing uncertainty: The choice of electoral formulas. *Public Choice, 141*(3–4), 391–403.

Karp, R. M. (1972). Reducibility among combinatorial problems. In R. E. Miller & J. W. Thatcher (Eds.), *Complexity of computer calculations* (pp. 85–103). New York: Plenum.

Kilgour, D. M. (2010). Using approval balloting in multi-winner elections. In Jean-François Laslier & M. R. Sanver (Eds.), *Handbook of approval voting* (pp. 105–124). Berlin: Springer.

Kilgour, D. M., & Marshall, E. (2011). Approval balloting for fixed-size committees. In M. Machover & D. Felsenthal (Eds.), *Electoral systems: Paradoxes, assumptions, and procedures* (pp. 305–326). Berlin: Springer.

Merrill, S., III, & Nagel, J. H. (1987). The effect of approval balloting on strategic voting under alternative decision rules. *American Political Science Review, 81*(2), 509–524.

Slinko, A., & White, S. (2010). Proportional representation and strategic voters. *Journal of Theoretical Politics, 22*(3), 301–332.

Still, J. W. (1979). A class of new methods for congressional apportionment. *SIAM Journal on Applied Mathematics, 37*(2), 401–418.

The Structure of Voters' Preferences Induced by the Dual Culture Condition

William V. Gehrlein and Souvik Roy

1 Introduction

A significant amount of research has been conducted in the past few decades to consider both the probability that various voting paradoxes might be observed and to evaluate common voting rules on the basis of their propensity to select desirable candidates. Much of the focus of this work has centered on elections with three candidates $\{A, B, C\}$ for n voters, where $A \succ B$ denotes that an individual voter prefers Candidate A to Candidate B. Individual voter indifference between candidates is not allowed, so that either $A \succ B$ or $B \succ A$ for all A and B. Intransitive voter preferences, such as $A \succ B$, $B \succ C$ and $C \succ A$, are prohibited as a requirement of individual rationality. There are therefore only six remaining possible complete preference rankings that each voter might have on the candidates, as shown in Fig. 1.

Here, n_i denotes the number of voters who have complete preferences on the candidates that are in agreement with the associated i^{th} preference ranking. For example there are n_3 voters with a preference ranking that has B being most preferred, C being least preferred and A being middle-ranked between the two.

Much of the work in this study was conducted while both authors were visiting at University of Caen.

W.V. Gehrlein (✉)
Department of Business Administration, University of Delaware, PO Box 178, Nassau, DE 19969 USA
e-mail: wvg@udel.edu

S. Roy
University of Caen, 14000 Caen, France

Economics and Planning Unit, Indian Statistical Institute, 7 S J S Sansanwal Marg, New Delhi 110 016, India
e-mail: souvik.2004@gmail.com

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare, DOI 10.1007/978-3-319-05158-1_19,

Fig. 1 Possible individual voter preference rankings on three candidates

A	A	B	C	B	C
B	C	A	A	C	B
C	B	C	B	A	A
n_1	n_2	n_3	n_4	n_5	n_6

Let $\boldsymbol{n}$ denote a six-dimensional vector, or voting situation, of such n_i terms for which $\sum_{i=1}^{6} n_i = n$.

We consider the comparison of candidates on the widely studied basis of Pairwise Majority Rule (PMR). Candidate A will beat Candidate B by PMR, which we denote as AMB, when more voters have $A \succ B$ than $B \succ A$. That is, AMB whenever $n_1 + n_2 + n_4 > n_3 + n_5 + n_6$. We assume throughout that n is odd to preclude the possibility of ties with PMR voting, and each voter is assumed to vote in PMR comparisons in accordance with their true preferences. Candidate A will be the PMR Winner, or Condorcet Winner, if both AMB and AMC. If some candidate is the Condorcet Winner in a given voting situation, then that candidate would certainly be deemed to be a good candidate to represent the overall most preferred candidate according to the preference rankings of the electorate.

It is widely known that a Condorcet Winner does not always exist, so that a voting situation can produce cyclical PMR relationships like AMB, BMC and CMA. Such an outcome is known as an occurrence of Condorcet's Paradox, when the PMR comparisons result in an intransitive relationship on pairs of candidates while the individual voters are prohibited from having such intransitivity. The Condorcet Criterion states that the Condorcet Winner should always be selected as the winner of an election whenever such a winner exists. Most commonly used voting rules cannot always meet the Condorcet Criterion; so in keeping with the intent of this criterion, the Condorcet Efficiency of a voting rule is the conditional probability that the voting rule will select the Condorcet Winner, given that such a winner exists.

Many analyses have been performed to determine factors that have an impact on both the probability that Condorcet's Paradox will be observed and the Condorcet Efficiency of voting rules. One frequent topic of consideration in these analyses has been the impact that social homogeneity might have on these two events. The term social homogeneity generally refers to the degree of dispersion that exists among voters' preferences. With larger relative measures of social homogeneity, voters' preferences will tend to become less disperse, or more alike, in nature. One would expect on an intuitive basis that as voters' preferences reflect greater degrees of social homogeneity with little dispersion, the likelihood of observing Condorcet's Paradox should decrease and that the Condorcet Efficiency of voting rules should increase. This is obviously true in the extreme case in which all voters have exactly the same preference ranking on candidates. However, it is quite surprising to note that very limited evidence has been provided to date to give direct support for either of these notions, despite a number of attempts that have been made to demonstrate these relationships. This current paper presents the first definitive study that clearly demonstrates the expected relationship between a classical measure of social homogeneity from the literature and both the probability that Condorcet's Paradox will be observed and the Condorcet Efficiency of voting rules.

In the next section we develop the notion behind a classic measure of social homogeneity and also present a classic model that has been used to describe the likelihood that various voting situations will be observed. Section 3 then considers a model that leads to the conclusion that the expected value of the probability of observing Condorcet's Paradox does indeed generally decrease as the specified measure of social homogeneity increases. Section 4 then shows that the expected Condorcet Efficiency of a frequently studied voting rule, known as Borda Rule, generally increases as the specified measure of social homogeneity increases. The final section then summarizes the conclusions of the study.

2 Social Homogeneity and the Dual Culture Condition

We begin this discussion by describing the likelihood that a voter who is selected at random from a population of voters will have a specified preference ranking on the three candidates, as shown in Fig. 2.

Here, p_i denotes the probability that a randomly selected voter will have complete preferences on the candidates that are in agreement with the i^{th} preference ranking, following the notion of the n_i terms in Fig. 1. Let $\boldsymbol{p}$ denote a six-dimensional vector of such p_i terms with $\sum_{i=1}^{6} p_i = 1$.

A random voting situation can then be obtained by sequentially drawing n random preference rankings to represent the voter's preferences, when the likelihood of drawing a specified ranking on each of the n draws is determined by $\boldsymbol{p}$. The sequential independent draws are done with replacement so that $\boldsymbol{p}$ remains fixed over each of the n draws. The total number of preference rankings of each type is then determined to produce the $\boldsymbol{n}$ for the associated random voting situation.

With this background, we are interested in a classic measure of social homogeneity that is given by $H(\boldsymbol{p})$, with:

$$H\left(\boldsymbol{p}\right) = \sum\nolimits_{i=1}^{6} p_i^2. \tag{1}$$

As noted above, the measure $H(\boldsymbol{p})$ gauges the amount of dispersion that is present among the p_i terms. It is maximized at $H(\boldsymbol{p}) = 1$ when $p_i = 1$ for some $1 \le i \le 6$, which corresponds to a completely homogeneous society in which every voter must have precisely the same preference ranking on the candidates. A Condorcet winner must obviously exist in this case, and any reasonable voting rule must then elect that Condorcet Winner. The measure $H(\boldsymbol{p})$ is minimized at $H(\boldsymbol{p}) = 1/6$ when $p_i = 1/6$ for each $1 \le i \le 6$, which corresponds to a completely non-homogeneous society in which every voter is equally likely to have any of the preference rankings on candidates. This specific case is denoted as the Impartial Culture Condition (IC) in the literature. For large electorates as $n \to \infty$, no candidate can be expected to

Fig. 2 Probabilities of possible individual voter preference rankings on three candidates

A	A	B	C	B	C
B	C	A	A	C	B
C	B	C	B	A	A
p_1	p_2	p_3	p_4	p_5	p_6

have any PMR advantage whatsoever over any other candidate with IC, to make the possible existence of a PMR cycle relatively easy to exist. Thus, the probability that Condorcet's Paradox will be observed should be exaggerated in this case. Since no candidate can be expected to have any advantage over any other candidate in the voters' preferences, it will also be difficult for voting rules to make any significant distinctions between the candidates. Thus, the Condorcet Efficiency of voting rules should tend to be low in this case. Our intuitive relationships from above therefore appear to be very reasonable assumptions at the two extreme points of the range of possible $H(\boldsymbol{p})$ values.

Abrams (1976) presented $H(\boldsymbol{p})$ as a measure of social homogeneity, and Gehrlein (2006) further refined this definition by specifying that $H(\boldsymbol{p})$ is a Population Specific Measure since it is based on the p_i parameters of the population that is used to generate random voting situations. A Situation Specific Measure of homogeneity would instead be based on the n_i values of a specified voting situation. Let $P_{Cycle}(3,n,\boldsymbol{p})$ denote the probability that a PMR cycle will be observed in a three-candidate election with n voters with the preferences of a randomly selected voter being described by $\boldsymbol{p}$. Abrams observed that the relationship between $H(\boldsymbol{p})$ and $P_{Cycle}(3,n,\boldsymbol{p})$ is not perfect. That is, it is possible to have $\boldsymbol{p}$ and $\boldsymbol{p}'$ such that $H(\boldsymbol{p}) > H(\boldsymbol{p}')$ while it is also true that $P_{Cycle}(3,n,\boldsymbol{p}) > P_{Cycle}(3,n,\boldsymbol{p}')$. But, while our intuitive relationship between $H(\boldsymbol{p})$ and $P_{Cycle}(3,n,\boldsymbol{p})$ might not always be valid in every single case, it might still be possible to show that this relationship can still be expected to be true in most cases.

Any observations about such an expectation will clearly be driven by the relative likelihood with which various $\boldsymbol{p}$ vectors will be observed. One common assumption of this type is the Dual Culture Condition (DC), which assumes that the set of feasible $\boldsymbol{p}$ vectors is such that each complete preference ranking on candidates is as likely to be drawn to represent a voter's preferences as its dual, or inverted ranking, with $p_1 = p_6$, $p_2 = p_5$, and $p_3 = p_4$. It is obvious with this assumption that $p_1 + p_2 + p_3 = p_4 + p_5 + p_6 = 1/2$. It is of interest to determine the behavior of $H(\boldsymbol{p})$ under the DC Assumption. This measure of social homogeneity will obviously be minimized with $H(\boldsymbol{p}) = 1/6$ for the case of IC, which is a special case of DC, following discussion above.

An extreme case that is very much in the spirit of DC is considered in a specific example that is proposed by Sen (1970) for a two-class society in which the classes have radically different interests. This "class war" condition can be expected to lead to voting situations that contain only two different voter preference rankings on candidates. One class would have some specified preference ranking and the other class would have the dual preference ranking. It is shown that PMR must always be transitive for odd n in this class war situation. If it is further assumed that the two classes in the population contain approximately the same number of members, Sen's model is equivalent to the $H(\boldsymbol{p})$ maximizing case for DC, which can easily be extended to:

Theorem 1. $P_{Cycle}(3, n, \boldsymbol{p}) = 0$ *for* $\boldsymbol{p}$ *in the DC subspace if any two of* $p_1, p_2, p_3 = 0$.

It therefore follows that $P_{Cycle}(3, n, \boldsymbol{p})$ is minimized when $H(\boldsymbol{p})$ is maximized for DC, in keeping with our intuitive relationships. Sen (1970) then goes on to significantly expand the class war model result with his famous condition of "extremal restrictions".

The DC assumption imposes a significant degree of structure on the preferences of voters in an electorate, by requiring a balance in preferences on pairs of candidates in which it is equally likely to have either $A \succ B$ or $B \succ A$ in any voter's preferences for all A and B. The impact of this requirement is discussed at length in Gehrlein and Lepelley (2012), where it is noted that this generally creates a scenario in which it should be relatively easy to produce a PMR cycle.

3 A Relationship Between $H(\boldsymbol{p})$ and $P_{Cycle}(3, n, \boldsymbol{p})$

Fishburn and Gehrlein (1980) considered the relationship between $H(\boldsymbol{p})$ and $P_{Cycle}(3, n, \boldsymbol{p})$, following earlier work by Niemi (1969), Jamison and Luce (1972), Fishburn (1973) and Kuga and Nagatani (1974). To begin describing their analysis, let $\boldsymbol{S}(\boldsymbol{p})$ denote the set of 6! different $s_i(\boldsymbol{p})$ vectors that correspond to the possible permutations of the six p_i terms within a given $\boldsymbol{p}$. Let $\overline{P}_{Cycle}(3, n, \boldsymbol{S}(\boldsymbol{p}))$ denote the average value of $P_{Cycle}(3, n, s_i(\boldsymbol{p}))$ probabilities, with

$$\overline{P}_{Cycle}(3, n, \boldsymbol{S}(\boldsymbol{p})) = \frac{\sum_{S(p)} P_{Cycle}(3, n, s_i(p))}{6!}. \tag{2}$$

This approach of using the average $P_{Cycle}(3, n, s_i(\boldsymbol{p}))$ over all permutations of $\boldsymbol{p}$, was used to deal with arguments that are presented in Abrams (1976) regarding the fact that values of $P_{Cycle}(3, n, \boldsymbol{p})$ could change dramatically as the p_i terms are interchanged within any $\boldsymbol{p}$, while it is obvious that it must simultaneously be true that $H(s_i(\boldsymbol{p})) = H(\boldsymbol{p})$ for each $s_i(\boldsymbol{p}) \in S(\boldsymbol{p})$.

Let $\Phi(\boldsymbol{p}|h)$ denote the subset of all possible $\boldsymbol{p}$ vectors for which $H(\boldsymbol{p}) = h$, for a specified h. A representation is obtained for the lower bound, $LB(3, n|h)$, of all possible $\overline{P}_{Cycle}(3, n, \boldsymbol{S}(\boldsymbol{p}))$ in $\Phi(\boldsymbol{p}|h)$ for given h. Then, a representation for the associated upper bound, $UB(3, n|h)$, is also obtained. It is shown that both $LB(3, 3|h)$ and $UB(3, 3|h)$ decrease as h increases for $n = 3$, to strongly suggest that there is a definite relationship between $H(\boldsymbol{p})$ and $\overline{P}_{Cycle}(3, 3, \boldsymbol{S}(\boldsymbol{p}))$ that is in agreement with our intuitive result. Unfortunately, this relationship was then found to become quite weak for large n, and our intuitive result is not found to hold at all when the space of possible $\boldsymbol{p}$ vectors is extended beyond DC to cover the space of all possible $\boldsymbol{p}$ with $\sum_{i=1}^{6} p_i = 1$.

The analysis in Fishburn and Gehrlein (1980) then goes on to develop a simple closed form representation for $P_{Cycle}(3,\infty,DC(\boldsymbol{p}))$, the limiting probability as $n \to \infty$ that a PMR cycle will exist for a $\boldsymbol{p}$ vector from the subspace of DC:

$$P_{Cycle}\,(3,\infty,DC(\boldsymbol{p})) = \frac{1}{4} - \frac{1}{2\pi}\sum\nolimits_{i=1}^{3} Sin^{-1}\,(1-4p_i). \tag{3}$$

It should be noted that this part of the study does not add the complication of the consideration of all permutations within $\boldsymbol{p}$ that was used in the immediately preceding discussion. An additional result is proved to show that there is some relationship between $H(\boldsymbol{p})$ and $P_{Cycle}(3,\infty,DC(\boldsymbol{p}))$ in the DC subspace. For any $\boldsymbol{p}$ in the DC subspace, assume without loss of generality that $p_1 > p_2$, and consider a $\boldsymbol{p}^*$ that is obtained by changing $p_1^* \to p_1+\delta$ and $p_2^* \to p_2-\delta$, while $p_3^* = p_3$. To keep $\boldsymbol{p}^*$ in the DC subspace p_i^* changes accordingly for $4 \le i \le 6$. It obviously follows that we must have $H(\boldsymbol{p}^*) > H(\boldsymbol{p})$ if $\delta > 0$, and it is further shown that $P_{Cycle}(3,\infty,DC(\boldsymbol{p}^*))$ will decrease as δ increases. Thus, the intuitive relationship between $H(\boldsymbol{p})$ and $P_{Cycle}(3,\infty,DC(\boldsymbol{p}))$ is directly observed for large electorates when $H(\boldsymbol{p})$ changes, if $\boldsymbol{p}$ is changed in this specified manner in the DC subspace. The specified requirement as to how $\boldsymbol{p}$ must be changed in order to obtain the observed result is however quite restrictive.

Gehrlein (1999) later provides a useful extension of Theorem 1 by showing that

Theorem 2. $P_{Cycle}(3,\infty,DC(\boldsymbol{p})) = 0$ *if any of* $p_1, p_2, p_3 = 0$.

The objective of this current study is to show that a much stronger relationship can be shown to exist between $H(\boldsymbol{p})$ and $P_{Cycle}(3,\infty,DC(\boldsymbol{p}))$ on an expected value basis, to give very strong evidence to support this intuitive relationship. We start by defining $\Phi(DC(\boldsymbol{p})|h)$ as the subset of all $\boldsymbol{p}$ vectors in the DC subset for which $H(\boldsymbol{p}) = h$.

Lemma 1. *For* $\boldsymbol{p} \in \Phi(DC(\boldsymbol{p})/h)$ *with* $1/6 \le h \le 1/2$ *and* $p_1 = p_6$, $p_2 = p_5$, *and* $p_3 = p_4$:

Either

$$p_2 = \frac{1}{4}\left\{1-2p_1+\sqrt{4p_1\,(1-3p_1)+4h-1}\right\} \tag{4}$$

and $$p_3 = \frac{1}{4}\left\{1-2p_1-\sqrt{4p_1\,(1-3p_1)+4h-1}\right\}.$$

Or

$$p_2 = \frac{1}{4}\left\{1-2p_1-\sqrt{4p_1\,(1-3p_1)+4h-1}\right\} \tag{5}$$

and $$p_3 = \frac{1}{4}\left\{1-2p_1+\sqrt{4p_1\,(1-3p_1)+4h-1}\right\}.$$

Proof. If $\boldsymbol{p} \in \Phi(DC(\boldsymbol{p})|h)$ then

$$p_1{}^2 + p_2{}^2 + \left(\frac{1}{2} - p_1 - p_2\right)^2 = \frac{h}{2}.$$

After expansion, this reduces to

$$2p_2{}^2 - 2\left(\frac{1}{2} - p_1\right)p_2 + \left[2p_1{}^2 - p_1 + \frac{1}{2}\left(\frac{1}{2} - h\right)\right] = 0.$$

The two possible values of p_2 above are derived directly from this result by using the solutions that are obtained from the quadratic equation with algebraic reduction. The two associated p_3 values are then obtained by using the DC relationship that $p_3 = \frac{1}{2} - p_1 - p_2$.□

The result of Lemma 1 tells us that $\Phi(DC(\boldsymbol{p})|h)$ can be obtained as the set of all pairs of vectors that are associated all feasible values of p_1 that can result in the specified value of h. The next logical step is to determine the ranges of feasible p_1 that are associated with a specified h. There are two restrictions that lead us to this result.

Lemma 2. *For $\boldsymbol{p} \in \Phi(DC(\boldsymbol{p})|h)$ with $p_1 = p_6$, $p_2 = p_5$, and $p_3 = p_4$:*

$$\frac{1 - \sqrt{12h - 2}}{6} \leq p_1 \leq \frac{1 + \sqrt{12h - 2}}{6}, \text{ for } 1/6 \leq h \leq 1/4 \tag{6}$$

$$0 \leq p_1 \leq \frac{1 + \sqrt{12h - 2}}{6}, \text{for } 1/4 \leq h \leq 1/2. \tag{7}$$

Proof. Each $\boldsymbol{p}$ must have a p_1 such that $H(\boldsymbol{p}) \leq h$ is feasible. Since $H(\boldsymbol{p})$ is minimized when $p_2 = p_3$ for any given p_1, it then follows for DC that it must be true that

$$p_1{}^2 + 2\left(\frac{\frac{1}{2} - p_1}{2}\right)^2 \leq \frac{h}{2}.$$

This relationship reduces to

$$\frac{3}{2}p_1{}^2 - \frac{1}{2}p_1 + \frac{1}{2}\left(\frac{1}{4} - h\right) \leq 0.$$

By taking the derivative of this function with respect to p_1 it is simple to show that it is uniquely minimized at $p_1 = 1/6$. The derivative also shows that the function is decreasing for $p_1 < 1/6$ and increasing for $p_1 > 1/6$. The range of feasible p_1 values for which $H(\boldsymbol{p}) \leq h$ in the statement of this lemma for $1/6 \leq h \leq 1/4$ is obtained by using the quadratic equation when this function is set at equality, so that $H(\boldsymbol{p}) - h = 0$. Moreover, both of these range endpoint values are within the feasible bounds of $0 \le p_1 \le 1/2$ for this range of h values. So, the entire region of p_1 between the specified endpoints where $H(\boldsymbol{p}) - h = 0$ has $H(\boldsymbol{p}) \le h$. The lower range endpoint for the feasible p_1 values is truncated at $p_1 = 0$ for $1/4 \le h \le 1/2$ since $\frac{(1-\sqrt{12h-2})}{6}$ is negative for $h < 1/4$. The upper range endpoint remains feasible over the region with $1/4 \le h \le 1/2$.□

Lemma 3. *For* $\boldsymbol{p} \in \Phi(DC(\boldsymbol{p})|h)$ *with* $p_1 = p_6$, $p_2 = p_5$, *and* $p_3 = p_4$*:*

$$p_1 \le \frac{1-\sqrt{4h-1}}{4} \text{ or } p_1 \ge \frac{1+\sqrt{4h-1}}{4}, \text{for } 1/4 \le h \le 1/2. \tag{8}$$

Proof. Each $\boldsymbol{p}$ must have a p_1 such that $H(\boldsymbol{p}) \ge h$ is feasible. Since $H(\boldsymbol{p})$ is maximized when $p_2 = \left(\frac{1}{2} - p_1\right)$ and $p_3 = 0$, or when $p_3 = \left(\frac{1}{2} - p_1\right)$ and $p_2 = 0$, for any given p_1, it then follows for DC that it must be true that

$$p_1{}^2 + \left(\frac{1}{2} - p_1\right)^2 \ge \frac{h}{2}.$$

This relationship reduces to

$$2p_1{}^2 - p_1 + \frac{1}{2}\left(\frac{1}{2} - h\right) \ge 0.$$

By taking the derivative of this function with respect to p_1 it is simple to show that it is uniquely minimized at $p_1 = 1/4$. The derivative also shows that the function is decreasing for $p_1 < 1/4$ and increasing for $p_1 > 1/4$. The p_1 values for which $H(\boldsymbol{p}) - h = 0$ in this function are obtained by using the quadratic equation when this function is set at equality, to find these two solution points at $p_1 = \frac{1 \pm \sqrt{4h-1}}{4}$. Given that this function is minimized at $p_1 = 1/4$, the shape of the function that was determined above leads directly to the statement of the lemma. The restriction that $1/4 \le h \le 1/2$ is needed to keep the two solution points from the quadratic equation to be real numbers within the feasible range with $0 \le p_1 < 1/2$.□

The combined results of Lemmas 1–3 can now be used to completely determine the set of all $\boldsymbol{p} \in \Phi(DC(\boldsymbol{p})|h)$. That is, Lemmas 2 and 3 give the ranges of feasible p_1 that can exist for $\boldsymbol{p} \in \Phi(DC(\boldsymbol{p})|h)$. And, for each of these feasible values of p_1 there are two $\boldsymbol{p}$ in $\Phi(DC(\boldsymbol{p})|h)$, with the pairs of values for p_2 and p_3 in these two $\boldsymbol{p}$ vectors being determined by Lemma 1. We make a technical note at this point that there will actually be only one $\boldsymbol{p}$ in $\Phi(DC(\boldsymbol{p})|h)$ if Lemma 1 results in $p_2 = p_3$ for a given p_1 and h. However, we will be assuming a uniform probability distribution over the range of all feasible p_1 when h has been specified in later analysis, so that the probability of this $p_2 = p_3$ outcome is of measure zero.

Lemma 2 gives the range, $\boldsymbol{R}^{-}(p_1|h)$, of feasible p_1 values when $1/6 \le h \le 1/4$, with

$$\boldsymbol{R}^{-}\left(p_1\middle|h\right) = \left\{\frac{1-\sqrt{12h-2}}{6} \le p_1 \le \frac{1+\sqrt{12h-2}}{6}\right\}. \tag{9}$$

The combined results of Lemmas 2 and 3 yield two distinct ranges of feasible p_1 values when $1/4 \le h \le 1/2$

$$\boldsymbol{R}_1^{+}\left(p_1\middle|h\right) = \left\{0 \le p_1 \le \frac{1-\sqrt{4h-1}}{4}\right\}, \text{ and} \tag{10}$$

$$\boldsymbol{R}_2^{+}\left(p_1\middle|h\right) = \left\{\frac{1+\sqrt{4h-1}}{4} \le p_1 \le \frac{1+\sqrt{12h-2}}{6}\right\}. \tag{11}$$

These results can now be used to obtain a representation for the limiting conditional expected value $E[P_{Cycle}(3,\infty,DC(\boldsymbol{p})|h)]$ of the probability that Condorcet's Paradox will be observed when it is assumed to be equally likely to observe all $\boldsymbol{p}$ in the DC subspace that have a given value of $H(\boldsymbol{p})=h$. This is accomplished by mimicking a procedure that evaluates expected values that are associated with voting models that dates back to de Laplace (1795).

The "total weighted sum" of $P_{Cycle}(3,\infty,DC(\boldsymbol{p}))$ values from Eq. (3) is obtained for the subspace of all $\boldsymbol{p}$ in DC for which $H(\boldsymbol{p})=h$, and this is denoted as $F_1(h)$. For the range of values $H(\boldsymbol{p})$ over $\boldsymbol{R}^{-}(p_1|h)$ with $1/6 \le h \le 1/4$, we assume from Eq. (9) that all $\frac{1-\sqrt{12h-2}}{6} \le p_1 \le \frac{1+\sqrt{12h-2}}{6}$ are equally likely to be observed. We know from Lemma 1 that there are two $\boldsymbol{p}$ vectors in the DC subspace that will give $H(\boldsymbol{p})=h$ for each of these p_1 values. Furthermore, Eqs. (4) and (5) show that these two $\boldsymbol{p}$ vectors are obtained by interchanging p_2 and p_3, and it is obvious that the representation for $P_{Cycle}(3,\infty,DC(\boldsymbol{p}))$ in Eq. (3) is invariant to permutations of the p_i terms in $\boldsymbol{p}$. So, both of the associated $\boldsymbol{p}$ vectors for a specified p_1 will yield the same value of $P_{Cycle}(3,\infty,DC(\boldsymbol{p}))$. The "total weighted sum" $F_1(h)$ is then obtained from

$$F_1(h) = 2\int_{\frac{1-\sqrt{12h-2}}{6}}^{\frac{1+\sqrt{12h-2}}{6}}\left[\frac{1}{4} - \frac{1}{2\pi}\left\{\begin{array}{c} Sin^{-1}(1-4p_1) + Sin^{-1}\left(2p_1 - \sqrt{4p_1(1-3p_1)+4h-1}\right) \\ + Sin^{-1}\left(2p_1 + \sqrt{4p_1(1-3p_1)+4h-1}\right)\end{array}\right\}\right]dp_1. \tag{12}$$

The "total sum" of possible $\boldsymbol{p}$ vectors in the DC subspace with $H(\boldsymbol{p})=h$ is denoted as $F_2(h)$ for the range of values $H(\boldsymbol{p})$ with $1/6 \le h \le 1/4$. With the

Fig. 3 Computed values of $E[P_{Cycle}(3, \infty, DC(\boldsymbol{p})|h)]$

h	$E[P_{Cycle}(3, \infty, DC(\boldsymbol{p})|h)]$
1/6	0.08774
.18	0.08416
.20	0.07787
.22	0.06984
.24	0.05822
.25	0.04746*
.26	0.05267
.28	0.05155
.30	0.04932
.32	0.04674
.34	0.04394
.36	0.04094
.38	0.03775
.40	0.03431
.42	0.03055
.44	0.02634
.46	0.02141
.48	0.01507
.50	0.00000

assumption that all $\frac{1-\sqrt{12h-2}}{6} \leq p_1 \leq \frac{1+\sqrt{12h-2}}{6}$ are equally likely to be observed, it then follows directly in the same fashion as our derivation of Eq. (12) that

$$F_2(h) = 2\int_{\frac{1-\sqrt{12h-2}}{6}}^{\frac{1+\sqrt{12h-2}}{6}} dp_1. \tag{13}$$

Then, a representation for $E[P_{Cycle}(3, \infty, DC(\boldsymbol{p})|h)]$ is obtained for $1/6 \leq h \leq 1/4$ as the ratio

$$E\left[P_{Cycle}\left(3, \infty, DC(\boldsymbol{p})\middle|h\right)\right] = F_1(h)/F_2(h). \tag{14}$$

While a closed form representation for $F_2(h)$ can be obtained trivially, there is no simple closed form representation for $F_1(h)$. This therefore leaves us with the option of obtaining values for $E[P_{Cycle}(3, \infty, DC(\boldsymbol{p})|h)]$ by using numerical methods. Computed values are listed for each h=.18(.02).24, along with $h = 1/6$ and $h = .25$, in Fig. 3.

These values show that $E[P_{Cycle}(3, \infty, DC(\boldsymbol{p})|h)]$ consistently decreases as h increases over $\boldsymbol{R}^-(p_1|h)$ with $1/6 \leq h \leq 1/4$, to strongly support our intuitive relationship between $H(\boldsymbol{p})$ and $P_{Cycle}(3, \infty, DC(\boldsymbol{p}))$.

A similar type of analysis can be used to obtain a representation for $E[P_{Cycle}(3, \infty, DC(\boldsymbol{p})|h)]$ for $1/4 \leq h \leq 1/2$. The only difference in this case is that the functions that are associated with $F_1(h)$ and $F_2(h)$ both will contain two integral components to account for the two regions of feasible p_1 values in $\boldsymbol{R}_1^+(p_1|h)$ and $\boldsymbol{R}_2^+(p_1|h)$. The resulting computed values are listed for each h=.26(.02).50 in Fig. 3. As noted in Theorem 1 and the discussion leading up to it, $P_{Cycle}(3,\infty,DC(\boldsymbol{p})|.50)=0$. These values show that $E[P_{Cycle}(3,\infty,DC(\boldsymbol{p})|h)]$ consistently decreases as h increases over $.26 \le h \le .50$, but there is one minor aberration that appears as a result of the fact that $E[P_{Cycle}(3,\infty,DC(\boldsymbol{p})|.25)] < E[P_{Cycle}(3,\infty,DC(\boldsymbol{p})|.26)]$.

There is clearly something unusual that happens in the neighborhood of $h=1/4$. As h increases from .25, Eqs. (10) and (11) show that $\boldsymbol{R}_1^+(p_1|h)$ and $\boldsymbol{R}_2^+(p_1|h)$ retract from their respective sides of the point at $p_1=1/4$ to drop $\boldsymbol{p}$ vectors with $p_1 \approx 1/4$ from the DC subspace with $H(\boldsymbol{p})=h$. It turns out that there is something very special about these $\boldsymbol{p}$ vectors that are dropped in the neighborhood of $h \approx 1/4$ when $p_1 \approx 1/4$. Suppose that $h \approx 1/4$ and $p_1 \approx 1/4$ for some $\boldsymbol{p}$ in the DC subspace. This can only be realized if either $p_2 \approx 1/4$ and $p_3=0$ or $p_2=0$ and $p_3 \approx 1/4$. Based on the result of Theorem 2 the probability of observing Condorcet's Paradox is zero for this particular $\boldsymbol{p}$. As a result, for small increases in h above .25 the $\boldsymbol{p}$ vectors that are being excluded as $\boldsymbol{R}_1^+(p_1|h)$ and $\boldsymbol{R}_2^+(p_1|h)$ retract will all have very small values for $P_{Cycle}(3,\infty,DC(\boldsymbol{p})|h)$, which results in the sudden and significant increase that is observed for $E[P_{Cycle}(3,\infty,DC(\boldsymbol{p})|h)]$ in moving from h=.25 to h=.26. It is important to note that this aberration only takes place for a very small region of $E[P_{Cycle}(3,\infty,DC(\boldsymbol{p})|h)]$ values over the entire possible range with $1/6 < h < 1/2$. Thus, we can indeed generally expect $E[P_{Cycle}(3,\infty,DC(\boldsymbol{p})|h)]$ to decrease as $H(\boldsymbol{p})$ increases with the assumption of DC.

Unfortunately, a number of follow-up studies failed to show any consistently significant relationship over the entire range of some population specific measures of social homogeneity, including $H(\boldsymbol{p})$, and $P_{Cycle}(3,n,\boldsymbol{p})$ when the space of possible $\boldsymbol{p}$ vectors is extended beyond DC to cover the space of all possible $\boldsymbol{p}$ with $\sum_{i=1}^{6} p_i = 1$. See for example: May (1971), Fishburn (1973), Gehrlein (1981) and Gehrlein (1987). The obvious conclusion that must be reached is that there is something that makes the DC subspace significantly different than the space of all possible $\boldsymbol{p}$ with $\sum_{i=1}^{6} p_i = 1$.

It turns out that $H(\boldsymbol{p})$ actually implies a much greater level of structure in voters' preferences, or group mutual coherence, than just social homogeneity when it is used in conjunction with DC. Suppose that $H(\boldsymbol{p})$ is maximized at $H(\boldsymbol{p})=1/2$ when $p_1=p_6=1/2$. Based on the preference rankings on candidates in Fig. 2, this means that every voter's preference ranking will have Candidate B as the middle-ranked preference. The same type of result follows for the other options that maximize $H(\boldsymbol{p})$, to require some candidate to be the middle-ranked candidate in every voter's preference ranking. Such a candidate is referred to as being a Perfectly Strong Centrist Candidate, since this candidate is neither strongly supported by nor strongly disliked by any of the voters in the electorate.

When $H(\boldsymbol{p})$ is minimized with IC as $n \to \infty$, there is expected to be an equal distribution of candidates over the middle-ranked preference position to reflect a scenario that is completely removed from having a Perfectly Strong Centrist Candidate. These results therefore suggest that $H(\boldsymbol{p})$ can be used as a population specific measure of the expected proximity of randomly generated voting situations to the condition of having a Perfectly Strong Centrist Candidate in the DC subspace of $\boldsymbol{p}$ vectors. As a result, $H(\boldsymbol{p})$ does more than just measure dispersion among voters' preferences with DC, it measures how close the population of voters is to having a very reasoned approach to determining the relative status of all candidates. We refer to this as representing group mutual coherence. However, this particular characteristic of $H(\boldsymbol{p})$ generally breaks down when attention is shifted from DC to the case in which all possible $\boldsymbol{p}$ vectors with $\sum_{i=1}^{6} p_i = 1$ can be observed.

4 A Relationship Between $H(\boldsymbol{p})$ and Condorcet Efficiency

The Condorcet Efficiency of a voting rule was defined previously as the conditional probability that the voting rule will select the Condorcet Winner, given that such a winner exists. As discussed above, intuition suggests that the Condorcet Efficiency of voting rules should generally tend to increase as voters' preferences become more homogeneous. This concept has been referred to as the Efficiency Hypothesis in the literature. We now examine the Efficiency Hypothesis for the limiting case of voters as $n \to \infty$ with DC for the case of a widely studied voting rule that is known as Borda Rule, by using $H(\boldsymbol{p})$ to measure the degree of homogeneity among voters' preferences.

Borda Rule is a special case of a general weighted scoring rule in which each voter reports their ranked preferences on the candidates and a different number of points is given to each candidate on the basis of its position in the voter's preference ranking. The winner of an election is then determined as the candidate that receives the greatest number of total points from voters. For the case of a three-candidate election, Borda Rule assigns one point to a candidate for each most preferred ranking in voters' preferences, one-half point for each middle position ranking in voters' preferences and zero points for each least preferred ranking in voters' preferences. Borda Rule has been shown uniquely to have a number of interesting properties.

Gehrlein (1999) develops a closed form representation for the Condorcet Efficiency of Borda Rule for three-candidate elections for the limiting case of voters as $n \to \infty$ with DC. This is denoted as $CE_{BR}(3, \infty, DC(\boldsymbol{p}))$, with

$$CE_{BR}(3, \infty, DC(\boldsymbol{p})) = \frac{\sum_{(i,j,k) \in \left\{ \begin{matrix} (1,2,3) \\ (2,1,3) \\ (2,1,3) \end{matrix} \right\}} \left[\frac{1}{2\pi} Sin^{-1}\left(\frac{1+2p_i}{\sqrt{6p_i+1}} \right) + \frac{1}{4\pi} \left\{ Sin^{-1}\left(\frac{2-2p_j-6p_k}{\sqrt{6p_i+1}} \right) + Sin^{-1}\left(\frac{2-2p_k-6p_j}{\sqrt{6p_i+1}} \right) \right\} \right]}{1 - P_{Cycle}(3, \infty, DC(\boldsymbol{p}))}. \tag{15}$$

Just as we observed in the representation for $P_{Cycle}(3, \infty, DC(\boldsymbol{p}))$ in Eq. (3), the representation for $CE_{BR}(3, \infty, DC(\boldsymbol{p}))$ in Eq. (15) is invariant to an interchange of p_2 and p_3.

It is also very simple to show that the Condorcet Winner that must exist when $H(\boldsymbol{p})$ is maximized at $H(\boldsymbol{p}) = 1/2$ under DC when either $p_1 = p_6 = 1/2$,

Fig. 4 Computed values of $E[CE_{BR}(3,\infty,DC(\boldsymbol{p})|h)]$

h	$E[CE_{BR}(3,\infty,DC(\boldsymbol{p})|h)]$
1/6	0.90119
.18	0.90229
.20	0.90407
.22	0.90600
.24	0.90784
.25	0.90764*
.26	0.91005
.28	0.91320
.30	0.91648
.32	0.92001
.34	0.92386
.36	0.92810
.38	0.93279
.40	0.93806
.42	0.94408
.44	0.95113
.46	0.95976
.48	0.97133
.50	1.00000

or $p_2 = p_5 = 1/2$, or $p_3 = p_4 = 1/2$ must also be elected as the winner by Borda Rule when n is odd. As mentioned previously, no voting rule can be expected to distinguish between candidates effectively when $H(\boldsymbol{p})$ is minimized with IC. So the Efficiency Hypothesis seems to be very valid at the extreme points of $H(\boldsymbol{p})$ when the DC assumption is used.

We use the representation in Eq. (15) and directly follow the notions of the preceding section that considered the relationship between $H(\boldsymbol{p})$ and $P_{Cycle}(3,\infty,DC(\boldsymbol{p}))$. We let $E[CE_{BR}(3,\infty,DC(\boldsymbol{p})|h)]$ denote the expected value of $CE_{BR}(3,\infty,DC(\boldsymbol{p}))$ when all $\boldsymbol{p}$ vectors in the DC subspace with $H(\boldsymbol{p}) = h$ are equally likely to be observed. A rather complex representation for $E[CE_{BR}(3,\infty,DC(\boldsymbol{p})|h)]$ is then obtained by mirroring the arguments that led to the development of $F_1(h)$ and $F_2(h)$ that led to $E[P_{Cycle}(3,\infty,DC(\boldsymbol{p})|h)]$, while substituting $CE_{BR}(3,\infty,DC(\boldsymbol{p}))$ for $P_{Cycle}(3,\infty,DC(\boldsymbol{p}))$. Resulting values for $E[CE_{BR}(3,\infty,DC(\boldsymbol{p})|h)]$ are obtained by using numerical methods, and these computed values are listed in Fig. 4 for each h=.18(.02).50, along with $h = 1/6$ and $h = .25$.

The results in Fig. 4 clearly indicate that the Efficiency Hypothesis is valid for Borda Rule over almost the entire range of $H(\boldsymbol{p})$ values in the DC subspace. The exception is that, just as in the case with $E[P_{Cycle}(3,\infty,DC(\boldsymbol{p})|h)]$, a small aberration occurs in the neighborhood of $h = 1/4$.

5 Conclusion

We know that it is not possible to show that a perfect relationship exists such that $P_{Cycle}(3,\infty,DC(\boldsymbol{p}))$ will always decrease as $H(\boldsymbol{p})$ increases for $\boldsymbol{p}$ in the DC subspace. However, we have been able to show that this intuitive relationship does indeed hold on an expected value basis over almost the entirety of the range of possible values of $H(\boldsymbol{p})$, and the aberration in the small region in which the relationship is reversed is completely explainable. We have also been able to show that the intuitive result that $CE_{BR}(3, \infty, DC(\boldsymbol{p}))$ increases as $H(\boldsymbol{p})$ increases for $\boldsymbol{p}$ in the DC subspace is also valid on an expected value basis over almost the entirety of the range of possible $H(\boldsymbol{p})$ values.

It has also been noted that these intuitive relationships tend to break down for Population Specific Measures of Social Homogeneity like $H(\boldsymbol{p})$ when attention is shifted from DC to the case in which all possible $\boldsymbol{p}$ vectors with $\sum_{i=1}^{6} p_i = 1$ can be observed. This is explained in our analysis by noting that $H(\boldsymbol{p})$ can be used as a population specific measure of the expected proximity of randomly generated voting situations to the condition of having a Perfectly Strong Centrist Candidate in the DC subspace of $\boldsymbol{p}$ vectors. As a result, $H(\boldsymbol{p})$ does more than simply measure the degree of homogeneity in voters' preferences with DC, it actually measures how close the population of voters is to having a very reasoned approach to determining the relative status of all of the candidates, which implies a measure of group mutual coherence. It can therefore be concluded that the presence of some measure of group mutual coherence that goes beyond simply measuring social homogeneity is a significant factor in the observation of the proposed intuitive results on an expected value basis.

References

Abrams, R. (1976). The voter's paradox and the homogeneity of individual preference orders. *Public Choice, 26*, 19–27.

Fishburn, P. C. (1973). Voter concordance, simple majority, and group decision methods. *Behavioral Science, 18*, 364–376.

Fishburn, P. C., & Gehrlein, W. V. (1980). Social homogeneity and Condorcet's paradox. *Public Choice, 35*, 403–420.

Gehrlein W.V. (1981). Non-profile specific measures of social homogeneity. Proceedings of Southeast Decision Sciences Institute (pp. 167–169). Charlotte, NC.

Gehrlein, W. V. (1987). A comparative analysis of measures of social homogeneity. *Quality and Quantity, 21*, 219–231.

Gehrlein, W. V. (1999). The Condorcet efficiency of Borda Rule under the dual culture condition. *Social Science Research, 28*(1999), 36–44.

Gehrlein, W. V. (2006). *Condorcet's paradox*. Berlin: Springer.

Gehrlein, W. V., & Lepelley, D. (2012). The value of research based on simple assumptions about voters' preferences. In D. S. Felsenthal & M. Machover (Eds.), *Electoral systems: Paradoxes, assumptions and procedures* (pp. 173–199). Berlin: Springer.

Jamison, D., & Luce, E. (1972). Social homogeneity and the probability of intransitive majority rule. *Journal of Economic Theory, 5*, 79–87.

Kuga, K., & Nagatani, H. (1974). Voter antagonism and the paradox of voting. *Econometrica, 42*, 1045–1067.

de Laplace PS. (1795). Analytic theory of probabilities. In F. Sommerlad & I. McLean (1991, Eds.), *The political theory of Condorcet II* (pp. 282–287). Oxford: University of Oxford, Working Paper.

May, R. M. (1971). Some mathematical remarks on the paradox of voting. *Behavioral Science, 16*, 143–151.

Niemi, R. G. (1969). Majority decision-making under partial unidimensionality. *American Political Science Review, 63*, 488–497.

Sen, A. K. (1970). *Collective choice and social welfare*. San Francisco: Holden-Day.

Three Apportionment Problems, with Applications to the United Kingdom

Iain McLean

1 Introduction: Three Overlapping Problems of Apportionment

Generically, the problem of apportionment may be defined as the problem of assigning a vector of integer numbers to each of a number of entitled entities that comes as close as possible to giving each entity its proportionate share of representation. Within that, there are a number of sub-problems. The correct solution (if a uniquely best solution exists) to one may not be the correct solution to another.

Case 1 A territory is divided into subnational units whose boundaries may not be crossed. An integer number of representatives must be assigned to each subnational unit.

This case has been widely discussed in practice since 1790, and the theoretical literature has now reached a clear conclusion. Most of the discussion has concerned the United States. The US Constitution requires: (US Const. I: 2)

> Representatives and direct Taxes shall be apportioned among the several States which may be included within this Union, according to their respective Numbers.... The actual Enumeration shall be made within three years after the first meeting of the Congress of the United States, and within every subsequent Term of ten Years, in such manner as they

An earlier version of this chapter was delivered at the "Workshop on Electoral Methods", Department of Mathematics, KTH, Stockholm, May 2011. Many thanks for comments received then especially from Svante Linusson, Svante Janson, and Paul Edelman. The usual disclaimer applies.

I. McLean (✉)
Nuffield College, University of Oxford, Oxford, UK
e-mail: iain.mclean@nu_eld.ox.ac.uk

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare,
DOI 10.1007/978-3-319-05158-1_20,

> shall by Law direct. The Number of Representatives shall not exceed one for every thirty thousand, but each State shall have at least one Representative.

Here the house is the US House of Representatives; the inviolable units are the States; and the constraints are those laid out in the constitutional text. The very first apportionment, in 1791, led to a conflict between the two leading Cabinet members of President Washington, namely Thomas Jefferson (from Virginia) and Alexander Hamilton (from New York).

Jefferson proposed a method which involved choosing a common divisor, dividing it into the representative population of each state, to obtain a quotient for each. All divisor methods are identical up to this point. They differ only in their treatment of the fractions or remainders, after the decimal point in the quotients. Jefferson's method was to drop all remainders. Call this the Jefferson method, one of the class of ***divisor*** methods.

Hamilton proposed a method which began by fixing total House size, and then dividing the qualifying population of each state into this number in order to obtain the exact theoretical entitlement, or quota, of representatives for each state.[1] Typically, and in the given case in 1791, to round all these quotas down would fill too few seats. To round them all up would fill too many seats. To round them off at the arithmetic mean, or any other fixed point, would sometimes assign too few seats and sometimes too many, Therefore a rounding procedure is required. Hamilton's method was first to award each state the number of seats in the lower integer bound of its quota. The surplus seats were then awarded to states in declining order of the remainder of their quotas until all seats had been allocated, and then the procedure stopped. Call this the ***Hamilton*** or ***largest-remainder*** method, one of the class of ***quota*** methods.

The constitutional requirement that each state must have at least one Representative introduces a further constraint, which was not binding in 1791, because for all House sizes proposed, each state's quota exceeded one. In subsequent apportionments the constraint has become closer to binding. For instance, in the apportionment of 2000, for a house size of 435, four states (ND, AK, VT, and WY) had quotas below one. The smallest (WY) had a quota of 0.766. As it happens, all the plausible quota and divisor methods[2] allocated Wyoming one seat in the given problem. If, however, a method assigned zero, that unit would have to be assigned one seat, taken out of the calculations, and the quota or divisor recalculated over the remainder of the house and the remainder of the subnational units. That would be required if the Jefferson method were used.

---

[1] 'Theoretical entitlement' is the phrase used by the Parliamentary Boundary Commission for England when it does these sums. In the theory of apportionment, the generic term is 'quota'. *Quota* must be carefully distinguished from the *quotients* used in divisor methods.

[2] Presumably with the exception of the Jefferson method. Although Balinski and Young (2001), Appendix B, shows a Jefferson apportionment of 1 for WY in 2000, this is presumably achieved by the assignment and recalculation discussed in the text.

The politics in 1791 were fierce and (of course) zero-sum; and they have remained so. Hamilton and his allies realised that Jefferson's method was biased towards large states, such as Virginia. Under Jefferson's method, Virginia got more than the upper integer bound of its quota. Jefferson and his allies retorted that Hamilton's method did not use a common divisor, as the Constitution required. The issue was settled in favour of Jefferson by President Washington (from Virginia) after taking advice from Edmund Randolph (from Virginia).

Over the nineteenth century other methods were proposed. Most relevantly, John Quincy Adams (from New England, a region of small states) proposed the mirror image of Jefferson's method: find a common divisor, and then award each state the integer above its quotient. As it was symmetrical with Jefferson's method, it systematically rewarded small states. James Dean (from Vermont, a small state) proposed rounding off at the harmonic mean: this is slightly biased towards the small. Daniel Webster (from Massachusetts, a medium-sized state) proposed rounding off at the arithmetic mean. Joseph A. Hill, a statistician at the Bureau of the Census, later proposed rounding off at the geometric mean (which has a bias to the small, but less so than the Dean method). The methods of Webster and Hill were taken up by two prominent academics, Walter F. Willcox and Edward V. Huntington. Huntington, the more eminent, persuaded the National Academy of Science to recommend Hill's method, which remains the method in use today.

The more intuitive quota methods remained in play until C. W. Seaton, chief clerk of the Census Office, noticed in 1880 that

> While making these calculations I met with the so-called "Alabama" paradox where Alabama was allotted 8 Representatives out of a total of 299, receiving but 7 when the total became 300. Such a result as this is to me conclusive proof that the process employed in obtaining it is defective, and that it does not in fact "apportion Representatives among the States according to their respective numbers" (*Congressional Record*, 47th Congress, 1st Session 1881, 12:704–5. Cited by Balinski and Young 2001: 38).

Seaton was right. His discovery is a deadly blow to all quota methods, not only Hamilton's. All of them are subject to several paradoxes of monotonicity, of which the Alabama Paradox is but one.

Balinski and Young (2001) have solved the apportionment problem for Case 1. They have proved that

- there is no method that avoids the population paradox and always stays within quota (p. 79)

where "stays within quota" means "never assigns a number other than the upper or lower integer bound of the quota". So there is an ineradicable choice between quota methods (which violate monotonicity) and divisor methods (which violate quota). However, some divisor methods almost always violate quota (e.g., the Jefferson and Adams methods) and others rarely do (e.g., Hill and Webster).

If one accepts the 1880 argument of the chief clerk of the US Bureau of the Census that monotonicity violations are always unacceptable, then one is left with only the set of divisor methods.

In a typical problem in Case 1, the rules specify that there must be no bias either to the large or to the small. In the US case, there is a potential bias to the small in the rule that each state must have at least one seat, but this is a condition of any representation at all for a small inviolable unit. It has also never arisen in practice. Apart from that, the Constitution mandates equal apportionment. This means it should contain no bias either to the small or to the large. Hence:

> [t]he only divisor method that is perfectly unbiased – which fully satisfies the principle of one-man, one-vote – is Webster's (Balinski and Young 2001: 85).

Case 2 Members are elected in a multi-member district by a list system of proportional representation. Given a vector of votes for each party, an integer number of seats must be assigned to each party that qualifies for any seat.

The mathematical structure of this problem is identical to that of Case 1. Policy issues may introduce differences, however. In Case 1, the rule-maker may wish to impose a threshold below which one unit is **guaranteed** a seat, as in the US Constitution. In Case 2, the rule-maker may wish to impose a threshold below which one unit is **denied** a seat, e.g., if it is a policy goal to reduce the fragmentation of parties in the legislature.

Perhaps the most remarkable part of Balinski and Young's (2001) work is that they were the first to see, after over two centuries of argument in the USA over Case 1, and 150 years of argument in regimes using list PR systems over Case 2, that the problems were one and the same. Not only that, but the algorithms invented in the USA to solve Case 1 were actually identical to algorithms invented in Europe and Latin America, to solve Case 2. The methods of implementation, as well as the context, were utterly different, but the math was exactly the same:

- The Jefferson method of apportionment is exactly the d'Hondt (dH) method of proportional representation;
- The Webster method of apportionment is exactly the Sainte-Laguë (S-L) method[3] of proportional representation; and
- The Hamilton method of apportionment is exactly the largest-remainder method of proportional representation.

Faced with this remarkable homology, the policy-maker has one easy task and one more difficult one. The easy task is to say that, for the same reason as in Case 1, any largest-remainder method should be discarded as being non-monotonic. The more difficult task is to recommend a system.

One argument is to say that there is no reason to be biased between the large and the small. This leads, analogously to Case 1, to a recommendation that unmodified S-L (Webster) should always be used.

[3]That is, the pure S-L method, in which the first divisor is 1, not the *ad hoc* modification used in Sweden and elsewhere, in which the first divisor is 1.4. The effect of this modification is somewhat to favour large parties.

Another argument is to say that there is some reason for a bias in favour of the large. In practice almost all known implementations for assigning seats to parties are biased to the large. They either use dH (Jefferson) or they modify S-L in a way that protects large parties. A very simple reason for this is that electoral rules are always decided by political parties except in the rare case that they are written into the constitution of a new country (e.g., USA 1787; Ireland 1921). The political parties in the legislature are usually large parties. They are large parties elected under a choice rule that favours large parties. Therefore, it is only natural that they tend to write electoral rules that are biased in favour of large parties. At the subsequent elections, the legislature is again composed of large parties, who have no incentive to change the electoral rule.

However, there is an argument of principle that is valid even when purveyed by merchants of self-interest. This is the excessive fractionalization argument mentioned above. How small is the smallest party to gain representation? This is essentially a function of district magnitude M. In any PR system in which V votes have been cast, the threshold for representation is the *Droop quota* $V/(M+1)$. Obviously, therefore, the larger the district magnitude, the smaller will be the smallest party that gains representation. Policy-makers who care about avoiding excessive fragmentation should therefore manipulate district magnitude as the primary means of avoiding it.

But there is a secondary issue that will be important in the UK case study to be presented later in this chapter. Most PR apportionments (including Single Transferable Vote and largest-remainder systems) will do roughly what dH (Jefferson) does, and award seats only to parties which achieve at least very close to the natural Droop quota.[4] But S-L (Webster) is different: it awards seats to parties that score roughly 0.5 times the natural (Droop) quota. A policymaker may wish to avoid S-L for the party apportionment problem, or to modify it *ad hoc*, as in Sweden. An alternative approach would be to impose a minimum threshold of representation, say of one Droop quota. This would then work in the rules like the requirement in the US Constitution for each state to have at least one Representative, but in the opposite direction. The calculation would again be similar for an open-list system (not for STV). At the first count of ballots, any party that had not achieved the threshold would be eliminated, and ballots for that party reassigned to the next valid place shown, then the S-L calculation would proceed.

Case 3 Seats in a supranational or intergovernmental body are to be assigned by shares of some relevant criterion/a: e.g., population, area, GDP per head, etc.

This problem has engaged a few mathematicians since the creation of the League of Nations, the United Nations and other supranational bodies (Penrose 1946; Richardson 1918, 1953). Penrose realised that there were good reasons not

[4]This statement is necessarily fuzzy because of complications caused by short ballots, i.e., ballots that do not express a full ranking of preferences among all the options.

to make representation proportionate to population. He proposed representation proportionate to the square root of population, or of electorate.

Penrose's model is stochastic: *The equitable representation of political opinion is partly a question of statistical probability* (Penrose 1946, p. 53). His set-up supposes that a substantial proportion of voters may be considered as randomly likely to vote either way on a binary proposition. In this model,

> The general formula for the probability of equal division of n random votes, where n is an even number, approaches $\sqrt{2/n\pi}$ when n is large. It follows that the power of the individual vote is inversely proportional to the square root of the number of people in the committee.

He continues by considering the case where there is a bloc of decisive votes and the rest may be presumed undecided (he is writing at the dawn of the United Nations and the emergence of the Soviet bloc):

> A bloc of size $\sqrt{n}$ always can carry 84 per cent. of the decisions, when the indifferent group has n voters, and a bloc of size $\sqrt[x]{n}$ can carry $2(1 + a)$ decisions where a is the area under the normal probability curve as given in the usual tables.

He concludes by saying that voting weights should be proportional to the square root, not of the population, but of the electorate, of member states. He believes (naively) that this will give a boost to democracies. After many years of neglect, the possibility of using the Penrose rule directly to assign voting weights in the Council of Ministers of the European Union has been floated recently (Slomczynski and Zyczkowski 2006).

Richardson (1953), a Quaker scientist who had been working on "Statistics of deadly quarrels" since World War I, wrote that for what became the League of Nations he had proposed

> that the assembly would probably deal only with affairs arising between nations, and would be prohibited by its constitution from interfering in affairs that are purely internal to a nation; and that therefore voting strength should be a measure of internationality. An index of internationality was suggested, having foreign trade as one of its ingredients (Richardson 1953: 697)

His proposal was rejected. The League of Nations was built on the "principle of the equality of sovereign states" (ibid.) on the principle of one state, one vote. So, by its Charter, Article 2, is the United Nations. Richardson noted that this principle was already tempered in 1953 by qualified-majority rules: 2/3 for important matters in the General Assembly, and the veto held by each of the five permanent members of the Security Council.

In essence, Penrose's insight is that voting weight is not the same as voting power. Richardson's insight is that the apportionment rule for seats on (or voting weight in) a supranational body ought not to be the same the apportionment rule for an intergovernmental body. How do these insights combine in the modern literature?

The literature on voting weight and voting power is now immense (for a summary see Felsenthal and Machover 1998). It suggests that for supranational bodies, voting weights should be assigned so that, as far as possible, each citizen should have equal voting power. Versions of the model may be used to propose apportionments of weights not only for the EU Council of Ministers, but also for bodies such as the IMF and the World Bank.

However, for intergovernmental bodies, the issues are different. A body which is truly intergovernmental can only function with the consent of all member governments. Therefore its voting rule is essentially the unanimity rule. Therefore apportionment is reduced to a trivial problem: no matter what apportionment of weights is used, the voting rule in practice is tantamount to “one state, one vote”.

In the next sections of this paper, I consider how these issues have been handled by UK policymakers since 1918. The focus is primarily on Case 1 and Case 2, but the lessons from Case 3—especially the lesson that voting weight is not the same as voting power—need to be borne in mind.

2 UK Approaches to the Problem of Apportioning Seats to Spatial Areas

UK policymakers’ use (or even understanding) of the Webster rule was limited or non-existent until 2004. Since then, in a remarkable transformation, it has become complete.

2.1 Assigning Seats in the House of Commons to the Four Countries of the UK (England; Scotland; Wales; Northern Ireland)

This section summarises a fuller account, with citations, in McLean (1995). Since 1832, the boundaries of the four countries of the UK have been regarded as inviolable. Initial allocations to Scotland (45) and to Ireland (100) were agreed in treaties of union between England and Scotland in 1707 and between Great Britain and Ireland in 1800. Neither was explicitly apportioned to population share: this was not an eighteenth-century way of looking at things. On entry to the successive unions, both Scotland and Ireland were under-represented in proportion to population. Wales was not administratively distinct, but its allocation at the start of modern apportionment, the 1832 Reform Act, was 31. Apportionment to Scotland and Ireland entered political debate in 1867–1868 and, more intensely, in 1884–1885, when there were franchise extension and redistribution packages (known to UK historians as the Second and Third Reform Acts). The Third Reform Act discussions marked the first attempts to discuss representation proportionate to population. In the apportionment in place from 1885 to 1918, Scotland was represented proportionately, as was Wales. Ireland had become over-represented not by an increase in its seats but by a decrease, both absolute and relative, in its population.

The apportionment of 1884–1885 was a one-off. It did not assign seats by formula, either within or between the countries of the UK, and unlike the US Constitution it contained no provisions for periodic reviews or automatic reapportionment. The next apportionment event therefore arose after the treaty of 1921 which created the Irish Free State, but left Northern Ireland as part of the UK. Northern Ireland was to have a devolved assembly. In the earlier, abortive, UK plan for devolution to the whole of Ireland, both Northern and southern Ireland would have remained in the UK, with reduced representation in the House of Commons, and Parliaments of Northern and Southern Ireland. What was left of that scheme after southern Irish independence was a Parliament of Northern Ireland, and a total of 12 territorial seats from Northern Ireland in the House of Commons. This was around 2/3 of its population share: deliberate under-representation, to acknowledge that the Parliament of Northern Ireland was now responsible for legislation on internal matters. But the number was by fiat, and again no provision for either reapportionment or automatic adjustment was made. A reapportionment for the rest of the UK in 1918 updated for the relative population figures within the countries of the UK but not between them, so that Scotland and Wales started to be overrepresented.

A new system was created in 1944. For the first time it allowed for periodic redistributions within each of the four countries, to track changing relative populations. But it did not tackle the between-country issue. Rather, it ossified it with the creation of four separate Boundary Commissions, one for each country. The minutes of the committee which created the system, not released until the 1990s, show that it was aware that both Scotland and Wales were overrepresented but that

> it would be very desirable, on political grounds, to state from the outset quite clearly that the number of Scottish and Welsh seats should not be diminished. The absence of any such assurance ... would lend support to the separatist movements in both countries (quoted in McLean 1995: 263).

It is a mystery that this commission discerned a separatist movement in Wales. Nobody else did, at that time.

Finally, the deliberate under-representation of Northern Ireland was unwound in a political deal which took effect in 1983. The Parliament of Northern Ireland had collapsed, and attempts to create a replacement had failed. Northern Ireland was directly ruled from Westminster and the rationale for deliberate under-representation had gone. The possibility was not raised when all three of the non-English territories gained parliaments or assemblies in 1997–1999.

Thus the 1944 regime, like its predecessors, failed to achieve the inter-country equity that was written into the US Constitution as a requirement for inter-state equity in 1787. Inter-country equity came about as a by-product of intra-country equity when the 1944 regime was swept away in 2010. To this we now turn.

2.2 Assigning Seats Within Each of the Four Countries of the UK

The 1944 regime was inherently conservative. The rules were written by incumbent politicians, and tended to protect the interests of incumbent politicians. Therefore, although electoral equality was one of the criteria to be observed in redistricting, it was not lexically prior to such other matters as respect for "local ties" and administrative boundaries. Unfortunately, nor was it lexically inferior. The rules, written by politicians, looped and recursed in such a way as to be mutually contradictory. This was pointed out, with increasing stridency, by a small number of academics (McLean and Mortimore 1992; McLean 1995, 2008; McLean and Butler 1996; Rossiter et al. 1999; Johnston et al. 2009).

The contradictions in the system arose because one rule (Rule 1) required the House of Commons not to continually increase in size; while at least three of the other rules had the effect, jointly and severally, of increasing its size at every redistricting. Of particular interest here is Rule 5, the equal-districting rule. Working within the normally-inviolable boundaries of administrative counties (Rule 4), Rule 5 stated,

> The electorate of any constituency shall be as near the electoral quota as is practicable having regard to the foregoing rules.

'Electoral quota' was defined as

> a number obtained by dividing the electorate for Great Britain by the number of constituencies in Great Britain existing on the enumeration date.

It might be thought—indeed the greatest C19 mathematician of voting, Lewis Carroll (C. L. Dodgson) thought—that expressing the equality rule, as here, in terms of mean constituency size, has the same effect as expressing it in terms of its reciprocal, the mean share of a legislator obtained by each elector. This is false. A few lines of algebra (McLean and Mortimore 1992) suffice to show that Rule 5 as written implied and was implied by the Dean (harmonic mean) apportionment rule. If Rule 5 had been written in the inverse fashion (each elector to have an equal share of an MP), it would have implied, and been implied by the Webster (arithmetic mean; S-L) rule. Rule 5 was therefore one of the rules which led to contradiction. As the harmonic mean is always below the arithmetic mean, fractional entitlements to seats would be (if Rule 5 were left to run unhindered) rounded up more often than down; therefore the House of Commons would get bigger at every redistricting. By the definition of "electoral quota" this enlargement would then be embedded as the starting point for the next review.

From the time of the Third Periodic Review of English constituency boundaries, which reported in 1983, the English Commissioners and their staff showed themselves to be aware of this anomaly. The issue arose whenever the theoretical entitlement of a county lay above the harmonic mean, but below the arithmetic mean, of two adjacent integers. In each of the Third, Fourth, and Fifth Reviews undertaken by the Boundary Commission for England, this case arose in a handful of counties. It was known internally as a “Walton” after the name of the Commissioner who spotted the anomaly. It always arises in the Isle of Wight, an island county with a theoretical entitlement that always lies above 1.33 (the harmonic mean of 1 and 2) and below 1.5. Rule 5 would always require the English Commission to give the Island two seats; but it always broke Rule 5 and awarded one. The Scottish Commission was even tougher, and held the Scottish seat total at 72 at the cost of violating both Rule 5 and Rule 4 (the rule about respecting local authority boundaries: Curtice 1996).

2.3 Assigning Seats in Other Legislatures (European Parliament; Proposed Senate; Scottish, Welsh, and Northern Ireland Assemblies/Parliaments)

Beginning with the first election to the European Parliament (EP) in 1979, the UK had used single-member districts in Great Britain, but always used the STV system to elect the three MEPs from Northern Ireland. Bowing to considerable pressure from the European Union, the UK switched to a list system of proportional representation for the GB seats with effect from the EP election of 1999. The districts used were, as they remain, the UK’s 12 standard statistical and administrative regions. Four of these (Scotland, Wales, Northern Ireland, and London) are political units of sub-national government; the other eight are not.

The enactment of what became the European Parliamentary Elections Act 1999 caused some political fun and games. The Minister responsible, Jack Straw (Labour) was not particularly interested in, or enthusiastic about, the EP. Introducing the bill he therefore cracked jokes about famous Belgians before introducing Victor d’Hondt. He stated that simulations of the previous EP election in 1994 conducted by his civil servants showed that dH produced fairer results than S-L for the party apportionment problem. Relevant extracts from his speech and later responses are in the Appendix.

Knowing *a priori* that the claim that dH would produce fairer results than S-L must be false, I asked Mr Straw’s department for the simulations. When they arrived some weeks later I searched for the error that they must contain, found it, and reported the results to the Minister and his department. Opposition politicians got to hear of this, with the results reported in the Appendix. Apart from the knockabout fun, the speeches and written answers in the Appendix reveal that for Case 1 (the apportionment of seats to regions), the UK government had used a largest-remainder (LR) method, although the minister’s statement does not mention the words *Hamilton* or *largest-remainder*. The LR apportionment of EP seats to the UK’s standard regions in 1999 was the same as an S-L apportionment would have been, but differed from the dH apportionment.

Opposition politicians tried to modify the use of dH for party apportionment in the 1999 bill, but failed, so the UK continues to use dH to apportion seats to parties for EP elections. However, as the next EP election in 2004 approached, civil servants realised that the UK needed a dynamic system for reapportioning its EP seats. The total number of seats might either rise or fall, as the number of member states and/or their relative populations fluctuated. Whether or not the civil servants realised it, the Hamilton method used in 1999 is not monotone with respect to population change and therefore it could not continue to be used without risking paradox. The task of reapportioning EP seats for 2004, in a way which would be robust for future reapportionments, fell to the statutory regulator, the Electoral Commission. The Commission issued a consultation paper outlining four possible ways of doing this. None of the four had any basis in the theory of apportionment. Although they had evidently talked to some academics before issuing the consultation, they had not located the UK-based scholars working in the field. Accordingly, a group of those, namely D.E. Butler, R. J. Johnston, A. McMillan, I. McLean, R. Mortimore, and H. P. Young, wrote to the Commission to point out that only S-L apportionment met the Commission's statutory criteria for apportionment of EP seats to regions. Unusually, the Commission discarded all four of its proposed methods, and adopted S-L (Electoral Commission 2003). They have continued to use S-L for their subsequent reapportionments (latest Electoral Commission 2010).

A search of the strings *hondt* and *saint* lague* in the UK Parliamentary Debates database (which covers the full text of debates and written answers in both houses of the UK Parliament for the years 1803–2005) was undertaken. No false positives were found: all references were to the two scholars and/or their divisor systems. There were 355 mentions of dH and 56 of S-L. The earliest was in 1966.[5] There was an intense concentration in 1998–1999 when not only the European Parliament bill but legislation affecting Scotland, Wales, and Northern Ireland was in hand. In 1998, following referendums in all three countries, legislation to create a Parliament in Scotland and Assemblies in Wales and Northern Ireland was enacted. All stipulated the use of PR.

The Northern Ireland task did not involve apportionment of seats, as that is done by STV for all internal elections there, but apportionment of ministerial portfolios. The current legislation for the Northern Ireland Assembly forces consociation. Previous assemblies were dominated by the majority unionist parties, which had no incentives to make any concessions to the large nationalist minority in Northern Ireland. The current regime first assigns the positions of First Minister and Deputy First Minister to the parties that come first and second in Assembly elections, and then assigns ministerial portfolios using dH. There has been no discussion in Parliament of using S-L for this task. It may be considered reasonable to use dH on the grounds that a bias towards large parties is justified in the interest of getting the business of the Assembly done.

Both Wales and Scotland were granted assemblies using mixed-member proportional systems. Each voter gets two votes, one to use in the district and one

[5] A slightly interesting negative finding is the lack of nineteenth-century references to d'Hondt. It seems that the UK's C19 debate about electoral systems was an insular affair.

for a party list. As in Germany and New Zealand, there are plurality elections in single-member districts—73 and Scotland and 40 in Wales. Those numbers are topped up to 129 (resp. 60) using the party list votes, so that as far as possible the party composition of the assembly represents the votes cast for the lists. There is no *überhangmandat* as in Germany where, if a party wins more district seats than its total list entitlement, the size of the assembly is increased temporarily to the extent of the overhang. In both countries the lists are regional, where each region is a superset of single-member districts.

In Scotland, the districts and regional boundaries were determined by fiat in a Schedule (Scotland Act 1998, Schedule 1). The districts were the parliamentary constituencies then existing; the regions were the European regional districts that were in process of being abolished in the European Parliament Elections Act going through the same Parliament. The only constraint on the size of regions appears to be

> The regional electorate of a region must be as near the regional electorate of each of the other regions as is practicable, having regard (where appropriate) to special geographical considerations. (Scotland Act 1998, Sch. 1, para. 13.

There is thus no explicit mechanism for regional reapportionment: neither Hamilton nor dH nor S-L. Likewise in Wales: the regions are defined by fiat. Although the Government of Wales Act 1998 was repealed and replaced in 2006, the provisions on districts and regions remained unchanged.

In contrast to the total lack of discussion about using an apportionment system to assign seats to regions, there was very full discussion in Parliament about which apportionment system to use for assigning seats to parties within each region. As with the European bill, both the Scottish and Welsh bills stipulated dH for the task. The parliamentary discussion was entirely partisan. Politicians in all parties knew that dH favoured large parties. Therefore, politicians from the large Labour and Conservative parties spoke out in favour of dH, and politicians from the small Liberal Democrat and nationalist parties spoke out in favour of S-L. No theoretical issues of interest were raised.

An analysis of the results of the 2009 European Parliament election in Great Britain showed that the results would have been quite different under S-L (Table 1).

Table 1 shows, as expected, that if the election had used S-L rather than dH apportionment of seats to parties, it would have produced a more proportionate result according to the standard Loosemore-Hanby index. The parties most affected would have been the Green Party and the BNP (British National Party). These findings produced a muted public reaction (except in the Green Party). Journalists who initially saw an interesting story backed off when they discovered that one of the beneficiaries of S-L would have been the far-right BNP. These results illustrate the trade-off between fairness and fractionalisation. No known apportionment system awards seats to nice small parties while withholding them from nasty small parties.

Table 1 Results of the 2009 election to the European Parliament: Great Britain

Party	Votes (%)	Actual seats (dH)	Seats under S-L	Actual % of seats	% of seats under S-L
Cons	27.7	25	21	36.2	30.4
UKIP	16.5	13	11	18.8	15.9
Labour	15.7	13	13	18.8	18.8
Lib Dem	13.7	11	11	15.9	15.9
Green	*8.6*	*2*	*7*	*2.9*	*10.1*
BNP	*6.2*	*2*	*3*	*2.9*	*4.3*
SNP	2.1	2	2	2.9	2.9
Plaid Cymru	0.8	1	1	1.5	1.4
Others	8.7	0	0	0.0	0.0
Total	100.0	69	69	100.0	100.0
Deviation from proportionality (Loosemore-Hanby 1971)				17.7	11.0

Source: adapted from McLean and Johnston (2009)

3 The Reapportionment Revolution in the UK

The strands in this paper came together when the coalition government elected in 2010 set out its constitutional programme. High up on the programme was a bill to equalise district size and to reduce the size of the House of Commons; and to maintain it thereafter at 600 members. Without doubt, a motive for this was the governing Conservatives' knowledge that existing system is biased against them: with equal votes and the current spatial distribution, the Labour Party wins many more seats than the Conservatives. In turn, one of the reasons for this is that Labour wins in small (and declining) seats, while the Conservatives win in large (and growing) seats. The system has been heavily lagged: for example, the 2010 General Election was fought on boundaries determined in 2000.

What has become the Parliamentary Voting System and Constituencies Act 2010 was therefore in the partisan interests of the Conservatives and against those of Labour, whose peers filibustered against it in the House of Lords in autumn–winter 2010. Since one of the ancestors of the British Labour Party is the Chartist movement, one of whose "Six Points" of 1848 was equal electoral districts, it was difficult for Labour to oppose the principle of equal districting, and they therefore focused on various secondary issues.

The bill offered an opportunity for civil servants and academics to come together in the common knowledge that the existing rules for the redistribution of seats were contradictory. Not only is it a bad idea (except for litigators) to enshrine contradictory rules in statute, but the contradictory nature of the old rules was one reason for the extreme delays in redistricting. Since any scheme must break at least one of the former statutory Rules, those aggrieved by it always had a good argument to hand, and protracted public inquiries were thus necessary.

In various publications (latest Johnston et al. 2009), the UK academics interested in apportionment had offered a template for a non-contradictory set of rules, in lexical order with electoral equality lexically prior to other criteria. This template was used as a reference point in drafting. After publication of the bill, a non-partisan "explainer" was published by the British Academy Policy Centre (Balinski et al. 2010) and sent to every legislator in the UK Parliament. The Bill survived the filibuster with few concessions to its critics (but one of them, ironically in view of past history, is to add the Isle of Wight to two other island areas already excluded from the equal-electorate requirement, and to offer it two seats, when Island lobbies were only seeking one).

In a small way this Act marks the culmination of a UK reapportionment revolution. The S-L (Webster) apportionment rule, which was unknown among policymakers until 1998 and then treated with suspicion or derision (although André Sainte-Laguë was not even a famous Belgian[6]), was adopted for EP elections in 2004 and has become uncontroversial in that application. It was adopted to assign seats to the countries of the UK in the 2010 Act, and was not one of the many controversial features of that Act. Whether or not it should be adopted for party apportionment remains open. As in other jurisdictions, the rules are made, mostly, by legislators from large parties, so it is likely that dH will remain the apportionment rule used for party apportionment in the UK.

Appendix. References to dH and S-L Apportionment in the UK Parliament

All cited from http://hansard.millbanksystems.com/

25.11.1997. Jack Straw [Labour minister (Home Secretary)]

Those who argue for Sainte-Lague say that it favours smaller parties and that d'Hondt discriminates in favour of larger parties. However, we do not believe that that is so, and we have calculated the likely effect of using all three divisors. The differences that they produce are minimal, and we have decided to use the d'Hondt divisor for four reasons. First, we believe that it will produce a fair result. Secondly, Sainte-Lague does not necessarily produce more proportional results. I have already introduced the proportionality index to the House—I noticed how hon. Members listened with bated breath. By calculating the index score for six regions using the 1994 European election vote in the United Kingdom, we found that, on average, d'Hondt scored higher than Sainte-Lague.

HC Deb 20 January 1998 vol 304 cc509-10W 509W

[6] André Sainte-Laguë (1882–1950) was a French mathematician educated at the Ecole Normale Supérieure, who became professor of applied mathematics at the Conservatoire National des Arts et Métiers; a left-wing activist and Résistance member during World War II.

§ Mr. Beith [Liberal Democrat]

To ask the Secretary of State for the Home Department, pursuant to his oral statement of 25 November 1997, Official Report, columns 812–13, whether he will publish the calculations on which his statement, regarding the proportionality of the Sainte-Lague and d'Hondt divisors, was based; if he will calculate the number of seats that would be won by each party in Scotland under the (a) d'Hondt and (b) Sainte-Lague divisors using the 1994 European Election results; and if he will make a statement.

§ Mr. George Howarth [Labour junior minister]

The calculations were based on the votes cast in the 1994 elections to the European Parliament in six of the 11 regions for which the European Parliamentary Elections Bill provides.

In five of the six regions the choice of divisor made no difference to the final allocation of seats. In Scotland, the effect was as follows:

d'Hondt:

- Four labour, one conservative,
- Three Scottish National Party

Sainte-Lague:

- Three labour, one conservative,
- One Scottish Liberal Democrat,
- Three Scottish National Party.

During the Bill's Second Reading debate, I gave the House figures which suggested that the two divisors produced different results in the London region.

Revised calculations show that both divisors produce the same result in London. I apologise for the original error.

§ Mr. Beith

To ask the Secretary of State for the Home Department what divisors were used to calculate the allocation of seats between the regions of England in the European Parliamentary Elections Bill. [24043]

§ *Mr. Straw*

[holding answer 19 January 1998]: No divisors were used. The allocation of seats to the English regions set out in Schedule 1 to the European Parliamentary Elections Bill was arrived at by dividing the total English electorate by the number of English seats (71) to produce an average figure. Seats were then allocated to regions in such a way as to ensure that the sum of the divergencies from this figure was as low as possible.

[James Clappison, (Conservative) 26.02.98]

Other people are more interested in this subject, and one of those is Professor Ian Maclean of Oxford university.

§ *Mr. Beith*

A Liberal Democrat.

§ *Mr. Clappison*

He may well be. He is certainly an expert on electoral systems, and he knows his stuff on these matters. The Home Secretary's comments came to his attention and he suspected that something was wrong—in fact, it was a bit more than a suspicion. It was impossible for the Home Secretary to be right, because the Sainte-Lague system is never less proportional than the d'Hondt system and is frequently more proportional than it—contrary to what the Home Secretary told us on Second Reading. The good professor thought that there had been a mistake—I am glad that the Home Secretary admits it now—he carried out simulations and he found that he was correct. I am glad that the Home Secretary has realised his mistake.

§ *Mr. Straw*

rose

§ *Mr. Clappison*

Before the Home Secretary intervenes again, I must advise him that it would have been a good idea for the Government to answer my written question on that subject.

§ Mr. Straw

rose—

§ The Chairman

Order. Mr. Clappison has the Floor.

§ Mr. Clappison

I shall certainly give way, but the Home Secretary might also like to respond to my next point. I asked the Government last week whether they had made a mistake, and I received the answer earlier this week that they were still thinking about it. The Home Secretary has obviously thought about it, so perhaps he can give us an answer.

§ Mr. Straw

The hon. Gentleman's implication is preposterous. He should know that, because I wrote to the shadow Home Secretary, his right hon. Friend the Member for North-West Cambridgeshire (Sir B. Mawhinney)—just as I wrote to the Liberal Democrat spokesman and other party leaders—as soon as we were aware that an error had been made. I apologised for that error and I have placed in the Library the details of the revised calculations. Furthermore, the Under—Secretary of State for the Home Department, my hon. Friend the Member for Knowsley, North and Sefton, East (Mr. Howarth), put on record in the Official Report of 20 January exactly what the situation was, and he repeated my unreserved apology to the House.

HC Deb 04 March 1998 vol 307 c627W

§ Mr. Clappison [Con]

To ask the Secretary of State for the Home Department what assessment he has made of the paper of 17 December 1997 written by Professor Ian McLean of Oxford University on the effect of different proportional representation formulae on the allocation of seats to parties in the European elections with particular reference to his conclusions on the use of the d'Hondt and Saint-Lague divisions.

§ Mr. George Howarth

[holding answer 23 February 1998]: Professor McLean's paper has been read with interest. The Government remain of the view that the d'Hondt divisor is the most suitable one to use for European Parliamentary elections, a view which was endorsed in a division on 26 February when the Committee of the whole House considered the European Parliamentary Elections Bill.

References

Balinski, M. L., Johnston, R., McLean, I., & Young, H. P. (2010). *Drawing a new constituency map for the United Kingdom: the parliamentary voting system and constituencies bill 2010*. London: British Academy Policy Centre.

Balinski, M. L., & Young, H. P. (2001). *Fair representation: Meeting the ideal of one man, one vote* (2nd ed.). Washington, DC: Brookings Institution Press.

Curtice, J. (1996). Should Britain follow the Caledonian way? In I. McLean & D. Butler (Eds.), *Fixing the boundaries: Defining and redefining single-member electoral districts* (pp. 235–250). Aldershot: Dartmouth.

Electoral Commission. (2003). *Distribution between electoral regions of UK MEPs.*http://www.electoralcommission.org.uk/__data/assets/pdf_file/0018/63018/DistributionofUKMEPsfinal_11167-8826__E__N__S__W__.pdf.

Electoral Commission. (2010). *Allocation of the additional MEP awarded to the UK under the treaty of Lisbon: recommendation.*http://www.electoralcommission.org.uk/__data/assets/pdf_file/0003/105960/MEP-Report-final-laid.pdf.

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

Johnston, R., McLean, I., Pattie, C., & Rossiter, D. (2009). Can the boundary commissions help the conservative party? Constituency size and electoral bias in the United Kingdom. *Political Quarterly, 80*(4), 479–494.

Loosemore, J., & Hanby, V. (1971). The theoretical limits of maximum distortion: Some analytic expressions for electoral systems. *British Journal of Political Science, 1*, 467–477.

McLean, I. (1995). Are Scotland and Wales overrepresented in the House of Commons? *Political Quarterly, 66*, 250–268.

McLean, I. (2008). Don't let the lawyers do the math: Some problems of legislative districting in the UK and the USA. *Mathematical and Computer Modelling, 48*, 1446–1454.

McLean, I., & Butler, D. E. (Eds.). (1996). *Fixing the boundaries: Defining and redefining single-member electoral districts*. Aldershot: Dartmouth.

McLean, I., & Johnston, R. (2009). When is proportional not proportional? Great Britain's 2009 elections to the European Parliament. *Representation, 45*, 4349–4355.

McLean, I., & Mortimore, R. G. M. (1992). Apportionment and the boundary commission for England (with). *Electoral Studies, 11*, 292–308.

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

Richardson, L.F. (various editors). (1993). *Collected papers of Lewis Fry Richardson* vol 2. 'International voting power', 1918, originally published in *War & Peace*, is at pp. 55–60. 'Voting in an international organization', unpublished MS, 1953, is at pp. 695–713.

Rossiter, D. J., Johnston, R. J., & Pattie, C. J. (1999). *The boundary commissions: Redrawing the UK's map of parliamentary constituencies*. Manchester: Manchester University Press.

Slomczynski, W., & Zyczkowski, K. (2006). Penrose voting system and optimal quota. *Acta Physica Polonica B, 37*, 3133–3143.

Part VI
List of Publications by Dan Felsenthal and Moshé Machover

The editors are grateful to Dan Felsenthal and Moshé Machover for providing this bibliography of their work to date. The following includes a comprehensive list of their jointly co-authored publications, followed by their works published individually and with other authors.

Dan Felsenthal and Moshé Machover: List of Joint Publications

A. Authored and Edited Books

1. Dan S. Felsenthal and Moshé Machover, *The Measurement of Voting Power: Theory and Practice, Problems and Paradoxes*. Cheltenham, UK, and Northampton, MA, USA: Edward Elgar Publishing Ltd., 1998, xviii + 322 pp.
2. Dan S. Felsenthal and Moshé Machover (eds.), *Electoral Systems: Paradoxes, Assumptions, and Procedures*. Berlin-Heidelberg: Springer, 2012, xii + 351 pp. ISBN 978-3-642-20440-1. eBook URL: http://dx.doi.org/10.1007/978-3-642-20441-8.

B. Articles in Refereed Journals

1. Dan S. Felsenthal and Moshé Machover, "After Two Centuries, Should Condorcet's Voting Procedure Be Implemented?", *Behavioral Science* 37:4 (October 1992), 250–274.
2. Dan S. Felsenthal and Moshé Machover, "Sequential Voting by Veto: Making the Mueller-Moulin Algorithm More Versatile", *Theory and Decision* 33:3 (November 1992), 223–240.
3. Dan S. Felsenthal and Moshé Machover, "Postulates and Paradoxes of Relative Voting Power—A Critical Re-Appraisal", *Theory and Decision* 38:2 (March 1995), 195–229.
4. Dan S. Felsenthal and Moshé Machover, "Who Ought to Be Elected and Who Is Actually Elected? An Empirical Investigation of 92 Elections Under Three Procedures", *Electoral Studies* 14:2 (June 1995), 143–169.
5. Dan Felsenthal and Moshé Machover, "Is This the Way to Elect a Prime Minister? The Pathological Properties of an Election Procedure To Be Used in Israel", (in Hebrew), *State, Government and International Relations* 40 (Summer 1995), 5–30.
6. Dan S. Felsenthal and Moshé Machover, "Alternative Forms of the Shapley Value and the Shapley-Shubik Index", *Public Choice* 87:3–4 (June 1996), 315–318.

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare, DOI 10.1007/978-3-319-05158-1_21,

7. Dan S. Felsenthal and Moshé Machover, "The Weighted Voting Rule in the EU's Council of Ministers, 1958–95: Intentions and Outcomes", *Electoral Studies* 16:1 (March 1997), 33–47.
8. Dan S. Felsenthal and Moshé Machover, "Ternary Voting Games", *International Journal of Game Theory* 26:3 (July 1997), 335–351.
9. Dan S. Felsenthal, Moshé Machover and William Zwicker, "The Bicameral Postulates and Indices of A Priori Relative Voting Power", *Theory and Decision* 44:1 (January 1998), 83–116.
10. Dan S. Felsenthal and Moshé Machover, "The Product Paradox of Voting Power", *Public Choice* 96:1–2 (July 1998), 81–92.
11. Dan S. Felsenthal and Moshé Machover, "Minimizing the Mean Majority Deficit: The Second Square-Root Rule", *Mathematical Social Sciences* 37:1 (January 1999), 25–37. Downloadable from http://eprints.lse.ac.uk/400/
12. Dan S. Felsenthal and Moshé Machover, "Myths and Meanings of Voting Power: Comments on a Symposium", *Journal of Theoretical Politics* 13:1 (January 2001), 81– 97. Downloadable also from http://eprints.lse.ac.uk/419/
13. Dan S. Felsenthal and Moshé Machover, "The Treaty of Nice and Qualified Majority Voting", *Social Choice and Welfare* 18:3 (July 2001), 431–464. Downloadable also from http://eprints.lse.ac.uk/420/
14. Dan S. Felsenthal and Moshé Machover, "Misreporting Rules", *Homo Oeconomicus* 17:4 (July 2001), 371–390. Downloadable from http://eprints.lse.ac.uk/24227/
15. Dan S. Felsenthal and Moshé Machover, "Annexations and Alliances: When Are Blocs Advantageous A Priori?", *Social Choice and Welfare* 19:2 (April 2002), 295–312. Downloadable also from http://eprints.lse.ac.uk/421/
16. Dan S. Felsenthal and Moshé Machover, "The Voting Power Approach: Response to a Philosophical Reproach", *European Union Politics* 4:4 (December 2003), 473–479, 493–497. Downloadable also from http://eprints.lse.ac.uk/422/
17. Dan S. Felsenthal and Moshé Machover, "A Priori Voting Power: What Is It All About?" *Political Studies Review* 2:1 (January 2004), 1–23. Downloadable also from http://eprints.lse.ac.uk/423/
18. Dan S. Felsenthal and Moshė Machover, "Analysis of QM Rules in the Draft Constitution for Europe Proposed by the European Convention, 2003", *Social Choice and Welfare* 23:1 (August 2004), 1–20. Downloadable also from http://eprints.lse.ac.uk/429/
19. Dan S. Felsenthal and Moshé Machover, "Qualified Majority Voting Explained", *Homo Oeconomicus* 21:3–4 (December 2004), pp. 573–595. Downloadable from http://eprints.lse.ac.uk/24230/
20. Dan S. Felsenthal and Moshé Machover, "Voting Power Measurement: A Story of Misreinvention", *Social Choice and Welfare* 25:2–3 (December 2005), pp. 485–506. Downloadable also from http://eprints.lse.ac.uk/550/ Reprinted in: Mark Haugaard & Stewart R. Clegg (eds.), *Power and Politics*, vol. 4, ch. 71, pp. 391–413. Sage Publications, 2012. ISBN: 978-0-85702-568-5

21. Dan S. Felsenthal and Moshé Machover, "The Majority Judgement Voting Procedure: A Critical Evaluation", *Homo Oeconomicus* 25:3/4 (December 2008), pp. 319–334. Downloadable from http://eprints.lse.ac.uk/24213/
22. Dan S. Felsenthal and Moshé Machover, "A Note on Measuring Voters' Responsibility", *Homo Oeconomicus* 26:2 (November 2009), pp. 259–271. Downloadable from http://eprints.lse.ac.uk/4296/
23. Dan S. Felsenthal and Moshé Machover, "The QM Rule in the Nice and Lisbon Treaties: Future Projections", in Manfred J. Holler and Mika T. Widgrén (eds.), *Essays in Honour of Hannu Nurmi* vol. I, *Homo Oeconomicus* 26:3–4 (December 2009), pp. 317–340. Downloadable from http://eprints.lse.ac.uk/2793/ Reprinted in: Manfred J. Holler and Hannu Nurmi (eds.), *Power, Voting, and Voting Power*: *30 Years After*. Berlin Heidelberg: Springer, 2013, Ch. 31, pp. 593–611.

C. Articles or Chapters in Books Which Are Not Conference Proceedings

1. Dan S. Felsenthal and Moshé Machover, "Models and Reality: The Curious Case of the Absent Abstention", in Manfred J. Holler and Guillermo Owen (eds.), *Power Indices and Coalition Formation*. Boston/Dordrecht/London: Kluwer Academic Publishers, 2001, pp. 87–103. Reprinted in *Homo Oeconomicus* 19:3 (December 2002), pp. 297–310, and in: Manfred J. Holler and Hannu Nurmi (eds.), *Power, Voting, and Voting Power*: *30 Years After*. Berlin Heidelberg: Springer, 2013, Ch. 4, pp. 73–86. Downloadable from: http://eprints.lse.ac.uk/archive/24228/
2. Dan S. Felsenthal and Moshé Machover, "Further Reflections on the Expediency and Stability of Alliances", in Matthew Braham and Frank Steffen (eds.), *Power, Freedom, and Voting*: *Essays in Honour of Manfred J. Holler*. Berlin Heidelberg: Springer, 2008, pp. 39– 55. Downloadable from: http://eprints.lse.ac.uk/archive/2566/

D. Articles in Conference Proceedings

1. Dan S. Felsenthal and Moshé Machover, "The Whole and the Sum of Its Parts: Formation of Blocs Revisited", in Manfred J. Holler, Hartmut Kliemt, Dieter Schmidtchen, and Manfred E. Streit (eds.), *Power and Fairness: Jahrbuch für Neue Politische Ökonomie 20*. Tübingen: Mohr Siebeck, 2002, pp. 279–291. (Proceedings of the NPÖ meeting on Power and Fairness, Bad Segeberg, Germany, 3–6 September, 2000.) Downloadable from http://eprints.lse.ac.uk/438/

E. Encyclopedia Entries

1. Dan S. Felsenthal and Moshé Machover composed the encyclopedia entries "I-power", "Penrose Voting Power Measure", "Power Indices", "P-Power", "Qualified Majority Voting", "Square-Root Rules". In Keith Dowding (ed.), *Encyclopedia of Power*. Thousand Oaks, CA: Sage, 2011, pp. 354–355, 476–477, 509–517, 528–529, 543–544, 627–628.

F. Reports

1. Dan S. Felsenthal and Moshé Machover, "The Power and Price of a Voter: A Note on the Shapley-Shubik and Banzhaf Indices", London School of Economics and Political Science, Centre for the Philosophy of the Natural and Social Sciences, Discussion Paper Series DP 2/94, July 1994.
2. Dan S. Felsenthal and Moshé Machover, "Ternary Voting Games", Hebrew University of Jerusalem, Center for the Study of Rationality and Interactive Decision Making, Discussion Paper No. 98, February 1996.
3. Dan S. Felsenthal and Moshé Machover, "Voting-Power and Parliamentary Defections: The 1953–54 French National Assembly Revisited". April 2000. Downloadable also from http://eprints.lse.ac.uk/594/
4. Dan S. Felsenthal and Moshé Machover, "Enlargement of the EU and Weighted Voting in its Council of Ministers", London School of Economics and Political Science, Centre for the Philosophy of the Natural and Social Sciences, Voting Power Report VPP 00/01, 16 November 2000. Downloadable also from http://eprints.lse.ac.uk/407/
5. Dan S. Felsenthal and Moshé Machover, "Myths and Meanings of Voting Power: Comments on a Symposium", London School of Economics and Political Science, Centre for the Philosophy of the Natural and Social Sciences, Voting Power Report VPP 01/01, January 2001. Downloadable also from http://eprints.lse.ac.uk/419/
6. Dan S. Felsenthal and Moshé Machover, "The Treaty of Nice and Qualified Majority Voting", London School of Economics and Political Science, Centre for the Philosophy of the Natural and Social Sciences, Voting Power Report VPP 01/02, Revised June 2001. Downloadable also from http://eprints.lse.ac.uk/420/
7. Dan S. Felsenthal and Moshé Machover, "Annexations and Alliances: When Are Blocs Advantageous A Priori?", London School of Economics and Political Science, Centre for the Philosophy of the Natural and Social Sciences, Voting Power Report VPP 02/01, April 2002. Downloadable also from http://eprints.lse.ac.uk/421/
8. Dan S. Felsenthal and Moshé Machover, "The Voting Power Approach: Response to a Philosophical Reproach", London School of Economics and Political Science, Centre for the Philosophy of the Natural and Social Sciences, Voting Power Report VPP 03/01, September 2003. Downloadable also from http://eprints.lse.ac.uk/422/
9. Dan S. Felsenthal and Moshé Machover, "A Priori Voting Power: What Is It All About?", London School of Economics and Political Science, Centre for the Philosophy of the Natural and Social Sciences, Voting Power Report VPP 04/05, January 2004. Downloadable also from http://eprints.lse.ac.uk/423/
10. Dan S. Felsenthal and Moshė Machover, "Analysis of QM Rules in the Draft Constitution for Europe Proposed by the European Convention, 2003", London School of Economics and Political Science, Centre for the Philosophy of the

Natural and Social Sciences, Voting Power Report VPP 04/03, August 2004. Downloadable also from http://eprints.lse.ac.uk/429/

11. Dan S. Felsenthal and Moshė Machover, "Population and Votes in the Council of Ministers: Squaring the Circle", Brussels, European Policy Centre, Issue Paper No. 10, (15.1.2004). This paper is Downloadable from: http://www.epc.eu/en/pub3.asp?TYP=TEWN&l=12&LV=187&see=y&PG=TEWN%2FEN%2Fart%5Fall%5Fpub&ll=13&offset=44
12. Dan S. Felsenthal and Moshė Machover, "Analysis of QM Rules Adopted by the EU Inter–Governmental Conference, Brussels, 18 June 2004". London School of Economics and Political Science, Centre for the Philosophy of the Natural and Social Sciences, Voting Power Report VPP 04/04, July 2004. Downloadable also from http://eprints.lse.ac.uk/431/
13. Dan S. Felsenthal and Moshé Machover, "Voting Power Measurement: A Story of Misreinvention", London School of Economics and Political Science, Centre for the Philosophy of the Natural and Social Sciences, Voting Power Report VPP 05/02, revised June 2005. Downloadable also from http://eprints.lse.ac.uk/550/
14. Dan S. Felsenthal and Moshé Machover, "Further Reflections on the Expediency and Stability of Alliances", London School of Economics and Political Science, Centre for the Philosophy of the Natural and Social Sciences, Voting Power Report VPP 06/01, revised August 2006. Downloadable from: http://eprints.lse.ac.uk/archive/2566/
15. Vincent C.H. Chua and Dan S. Felsenthal, "Coalition Formation Theories Revisited: An Empirical Investigation of Aumann's Hypothesis", London School of Economics and Political Science, Centre for the Philosophy of the Natural and Social Sciences, Voting Power Report VPP 06/xx, revised August 2006. Downloadable from: http://eprints.lse.ac.uk/767/
16. Dan S. Felsenthal and Moshé Machover, "Analysis of QM Rule Adopted by the Council of the European Union, Brussels, 23 June 2007", London School of Economics and Political Science, Centre for the Philosophy of the Natural and Social Sciences, Voting Power and Procedures Project, July 2007. Downloadable from http://eprints.lse.ac.uk/2531/
17. Dan S. Felsenthal and Moshé Machover, "The QM Rule in the Nice and EU Reform Treaties: Future Projections", London School of Economics and Political Science, Centre for the Philosophy of the Natural and Social Sciences, Voting Power and Procedures Project, October 2007. Downloadable from http://eprints.lse.ac.uk/2793/
18. Dan S. Felsenthal and Moshé Machover, "A Note on Measuring Voters' Responsibility", London School of Economics and Political Science, Centre for the Philosophy of the Natural and Social Sciences, Voting Power and Procedures Project, March 2008. Downloadable from http://eprints.lse.ac.uk/4296/

Dan Felsenthal: List of Publications

[Note: This list excludes joint publications with Moshé Machover]

A. Authored Books

1. Dan Felsenthal, *Mathematics for Administrative Decision Makers*, *Vol. I*, (in Hebrew). Tel-Aviv: Administrative Library, 1976, 251 pp.
2. Dan Felsenthal, *Mathematics for Administrative Decision Makers*, *Vol. II*, (in Hebrew). Tel-Aviv: Administrative Library, 1976, 267 pp.
3. Dan S. Felsenthal, *Topics in Social Choice: Sophisticated Voting, Efficacy, and Proportional Representation*. New York: Praeger, 1990, xx + 208 pp. ISBN 978-0-275-93430-9

B. Articles in Refereed Journals

1. Dan Felsenthal, "On the Relationship Between the Number of Ministries and Ministers and the Coalition Structure in Israel", (in Hebrew), *Netivey Irgun Ve'Minhal* 1–2: (May 1974), 26–34.
2. Robert J. Weiss, Joel C. Kleinman, Ursula C. Brandt and Dan S. Felsenthal, "The Effect of Importing Physicians: Return to a Pre-Flexnerian Standard", *The New England Journal of Medicine* 290:26 (June 1974), 1453–1458.
3. Joel C. Kleinman, Robert J. Weiss and Dan S. Felsenthal, "Physician Manpower Data: The Case of the Missing Foreign Medical Graduates", *Medical Care* 12:11 (November 1974), 906–917.
4. Joel C. Kleinman, Robert J. Weiss and Dan S. Felsenthal, "Immigrant Physicians: Results of a Cohort Study", *Inquiry* 13:3 (September 1975), 193–203.
5. Dan S. Felsenthal and Eliezer Fuchs, "An Experimental Evaluation of Five Designs of Redundant Organizational Systems", *Administrative Science Quarterly* 21 (September 1976), 474–488.
6. Dan S. Felsenthal, "Bargaining Behavior when Profits are Unequal and Losses are Equal", *Behavioral Science*, 22:5 (September 1977), 334–340.

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare,
DOI 10.1007/978-3-319-05158-1_22,

7. Abraham Diskin and Dan S. Felsenthal, "Decision Making in Mixed Situations in which Both Chance and a Rival Player are Confronted Simultaneously", *Behavioral Science* 23:4 (July 1978), 256–263.
8. Dan S. Felsenthal, "Aspects of Coalition Payoffs: The Case of Israel", *Comparative Political Studies* 12:2 (July 1979), 151–168.
9. Dan S. Felsenthal, "Group Versus Individual Gambling Behavior: Reexamination and Limitation", *Behavioral Science* 24:5 (September 1979), 334–345.
10. Dan Felsenthal, "Review of Some Aspects of Social Choice: Cyclical Majorities Versus Inefficient Voting Outcomes", (in Hebrew), *State, Government and International Relations* 15 (Spring 1980), 57–73.
11. Dan S. Felsenthal, "Applying the Redundancy Concept to Administrative Organizations", *Public Administration Review* 40 (May–June 1980), 247–252.
12. Dan S. Felsenthal and Abraham Diskin, "Decision Making in Mixed Situations: An Application to Israeli-Egyptian Relations, 1956–1979", *International Interactions* 7:1 (1980), 33–55.
13. Abraham Diskin and Dan S. Felsenthal, "Do They Lie?", *International Political Science Review* 2:4 (October 1981), 407–422.
14. Abraham Diskin and Dan S. Felsenthal, "An Experimental Evaluation of Samson's Dilemma", *Conflict Management and Peace Science* 5:2 (Spring 1981), 121–138.
15. Dan S. Felsenthal and Abraham Diskin, "Two Bargaining Solutions: An Experimental Reevaluation", *Simulation and Games* 13:2 (June 1982), 179–197.
16. Dan Felsenthal, "Strategic Considerations Under Cumulative Voting", (in Hebrew), *State, Government and International Relations* 19–20 (Spring 1982), 144–159.
17. Dan S. Felsenthal and Abraham Diskin, "The Bargaining Problem Revisited: Minimum Utility Point, Restricted Monotonicity Axiom, and the Mean as an Estimate of Expected Utility", *Journal of Conflict Resolution* 26:4 (December 1982), 664–691.
18. Dan S. Felsenthal, "Is Cumulative Voting Really Different from One-Man One-Vote?", *Electoral Studies* 4:2 (August 1985), 142–148.
19. Dan S. Felsenthal and Avraham Brichta, "Sincere and Strategic Voters: An Israeli Study", *Political Behavior* 7:4 (December 1985), 311–324.
20. Dan S. Felsenthal, Ze'ev Maoz and Amnon Rapoport, "Comparing Voting Systems in Genuine Elections: Approval-Plurality Versus Selection-Plurality", *Social Behaviour* 1:1 (September 1986), 41–53.
21. Ze'ev Maoz and Dan S. Felsenthal, "Self-Binding Commitments, the Inducement of Trust, Social Choice, and the Theory of International Cooperation", *International Studies Quarterly* 31:1 (March 1987), 177–200.
22. Amnon Rapoport, Dan S. Felsenthal and Ze'ev Maoz, "Microcosms and Macrocosms: Seat Allocation in Proportional Representation Systems", *Theory and Decision* 24:1 (January 1988), 11–33.

23. Dan S. Felsenthal and Ze'ev Maoz, "A Comparative Analysis of Sincere and Sophisticated Voting under the Plurality and Approval Procedures", *Behavioral Science* 33:2 (April 1988), 116–130.
24. Dan S. Felsenthal, Amnon Rapoport and Ze'ev Maoz, "Tacit Co-Operation in Three-Alternative Non-Cooperative Voting Games: A New Model of Sophisticated Behaviour under the Plurality Procedure", *Electoral Studies* 7:2 (August 1988), 143–161.
25. Amnon Rapoport, Dan S. Felsenthal and Ze'ev Maoz, "Proportional Representation: An Empirical Evaluation of Single-Stage, Non-Ranked Voting Procedures", *Public Choice* 59:2 (November 1988), 151–165.
26. Dan S. Felsenthal, "On Combining Approval with Disapproval Voting", *Behavioral Science* 34:1 (January 1989), 53–60.
27. Amnon Rapoport and Dan S. Felsenthal, "Efficacy in Small Electorates Under Plurality and Approval Voting", *Public Choice* 64:1 (January 1990), 57–71.
28. Dan S. Felsenthal, Ze'ev Maoz and Amnon Rapoport, "The Condorcet-Efficiency of Sophisticated Voting under the Plurality and Approval Procedures", *Behavioral Science* 35:1 (January 1990), 24–33.
29. Amnon Rapoport, Eythan Weg and Dan S. Felsenthal, "Effects of Fixed Costs in Two-Person Bargaining", *Theory and Decision* 28:1 (January 1990), 47–71.
30. Eythan Weg, Amnon Rapoport and Dan S. Felsenthal, "Two-Person Bargaining Behavior in Fixed Discounting Factor Games with Infinite Horizon", *Games and Economic Behavior* 2:1 (March 1990), 76–95.
31. Dan S. Felsenthal, "Averting the Quorum Paradox", *Behavioral Science* 36:1 (January 1991), 57–63.
32. Dan S. Felsenthal and Ze'ev Maoz, "Normative Criteria of Four Single-Stage Multi-Winner Electoral Procedures", *Behavioral Science* 37:2 (April 1992), 109– 127.
33. Dan S. Felsenthal, "Proportional Representation Under Three Voting Procedures: An Israeli Study", *Political Behavior* 14:2 (June 1992), 159–192.
34. Abraham Diskin, André Eschet-Schwarz, and Dan S. Felsenthal, "Homogeneity, Heterogeneity and Direct Democracy: The Case of Swiss Referenda", *Canadian Journal of Political Science* 40: 2 (June 2007), pp. 317–342. Downloadable from: http://eprints.lse.ac.uk/24231/
35. Paul R. Abramson, Abraham Diskin, and Dan S. Felsenthal, "Nonvoting and the Decisiveness of Electoral Outcomes", *Political Research Quarterly* 60:3 (September 2007) pp. 500–515. Downloadable from: http://eprints.lse.ac.uk/24232/
36. Abraham Diskin and Dan S. Felsenthal, "Individual Rationality and Bargaining", *Public Choice* 133: 1–2 (October 2007), pp. 25–29. Downloadable from: http://eprints.lse.ac.uk/24233/; from: http://www.springerlink.com/content/x27084q2h427370u/fulltext.pdf as well as from: http://dx.doi.org/10.1007/s11127-007-9212-7
37. Vincent C.H. Chua and Dan S. Felsenthal, "Do Voting Power Considerations Explain the Formation of Political Coalitions? A Re-Evaluation", *Homo Oeconomicus* 25:2 (November 2008) , pp. 141–167. Downloadable from

http://eprints.lse.ac.uk/24234/ Reprinted in: Manfred J. Holler and Hannu Nurmi (eds.), *Power, Voting, and Voting Power: 30 Years After*. Berlin Heidelberg: Springer, 2013, Ch. 23, pp. 443–466.

38. Dan S. Felsenthal and Nicolaus Tideman, "Varieties of Failure of Monotonicity and Participation under Five Voting Procedures", *Theory and Decision* 75:1 (July 2013), pp. 59–77. DOI 10.1007/s11238-012-9306-7. Downloadable from: http://www.springerlink.com/openurl.asp?genre=article&id=doi:10.1007/s11238-012-9306-7

C. Articles or Chapters in Books Which Are Not Conference Proceedings

1. Steven J. Brams, Dan S. Felsenthal and Ze'ev Maoz, "New Chairman Paradoxes", in Andreas Diekmann and Peter Mitter (eds.), *Paradoxical Effects of Social Behavior: Essays in Honor of Anatol Rapoport*. Heidelberg-Wien: Physica-Verlag, 1986, pp. 243–256.
2. Steven J. Brams, Dan S. Felsenthal and Ze'ev Maoz, "Chairman Paradoxes under Approval Voting", in Gerald L. Eberlein and Hal Berghel (eds.), *Theory and Decision: Essays in Honor of Werner Leinfellner*. Dordrecht: Reidel, 1988, pp. 223–233.
3. Amnon Rapoport, Dan S. Felsenthal and Ze'ev Maoz, "Sincere Versus Strategic Voting Behavior in Small Groups", in Thomas R. Palfrey (ed.), *Laboratory Research in Political Economy*. Ann Arbor: University of Michigan Press, 1991, pp. 201–235.
4. Vincent C.H. Chua and Dan S. Felsenthal, "Coalition Formation Theories Revisited: An Empirical Investigation of Aumann's Hypothesis", in Matthew Braham and Frank Steffen (eds.), *Power, Freedom, and Voting: Essays in Honour of Manfred J. Holler*. Berlin Heidelberg: Springer, 2008, pp. 159–183. Downloadable from: http://eprints.lse.ac.uk/767/
5. Dan S. Felsenthal, "Some Informal Remarks on Devising a 'Fair' Decision-Making Rule for Representative Assemblies", in Dan S. Felsenthal and Moshé Machover (eds.), *Electoral Systems: Paradoxes, Assumptions, and Procedures*. Berlin Heidelberg: Springer, 2012, Ch. 2, pp. 11–15.
6. Dan S. Felsenthal, "Review of Paradoxes Afflicting Procedures for Electing a Single Candidate", in Dan S. Felsenthal and Moshé Machover (eds.), *Electoral Systems: Paradoxes, Assumptions, and Procedures*. Berlin Heidelberg: Springer, 2012, ch. 3, pp. 19–91.

D. Articles in Conference Proceedings

1. Dan S. Felsenthal, "Priorities and Decision Making in Health Planning", *Israel Journal of Medical Sciences*, 10:1–2 (January–February 1974), 74–79. (Proceedings of the Batsheva Seminar on Health Administration and Health Economics, Beersheba, Israel, 27 December 1972 – 3 January 1973.)

E. Reports and Book Reviews

1. Dan Felsenthal, *The Policy of Israeli Governments Towards Higher Education.* (Hebrew). Master's Thesis, Hebrew University of Jerusalem, 1967.
2. Dan Felsenthal, *The Effects of Out-of-Hospital Physicians' Care Insurance Coverage on the Utilization of Hospitals in the United States: Study of a Health Policy-Making Problem*. PhD Thesis, Massachusetts Institute of Technology, 1971.
3. Dan S. Felsenthal, Ze'ev Maoz and Amnon Rapoport, "Comparing Approval with Plurality Voting in Genuine Elections", University of Haifa, Institute of Information Processing and Decision Making, Report No. 21, February 1985.
4. Amnon Rapoport, Dan S. Felsenthal and Ze'ev Maoz, "Microcosms and Macrocosms: Seat Allocation in Proportional Representation Systems", University of Haifa, Institute of Information Processing and Decision Making, Report No. 34, April 1986.
5. Amnon Rapoport, Dan S. Felsenthal and Ze'ev Maoz, "Proportional Representation in Israel's Federation of Labor: An Empirical Evaluation of a New Scheme", University of Haifa, Institute of Information Processing and Decision Making, Report No. 35, April 1986.
6. Amnon Rapoport, Dan S. Felsenthal and Ze'ev Maoz, "Sincere Versus Sophisticated Behavior in Noncooperative Four-Person Voting Games", University of Haifa, Institute of Information Processing and Decision Making, Report No. 38, June 1986.
7. Dan S. Felsenthal, Amnon Rapoport and Ze'ev Maoz, "Tacit Cooperation in Three- Alternative Noncooperative Voting Games: A New Model of Sophisticated Behavior Under the Plurality Procedure", University of Haifa, Institute of Information Processing and Decision Making, Report No. 47, June 1987.
8. Amnon Rapoport and Dan S. Felsenthal, "Efficacy in Small Electorates Under Plurality and Approval Voting", University of Haifa, Institute of Information Processing and Decision Making, Report No. 70, March 1989.
9. Amnon Rapoport, Eytan Weg and Dan S. Felsenthal, "Effects of Fixed Costs in Two-Person Sequential Bargaining", University of Haifa, Institute of Information Processing and Decision Making, Report No. 73, January 1990.
10. Amnon Rapoport, Eytan Weg and Dan S. Felsenthal, "Two-Person Bargaining Behavior in Fixed Discounting Factors Games with Infinite Horizon", University of Haifa, Institute of Information Processing and Decision Making, Report No. 74, January 1990.
11. Dan S. Felsenthal, Ze'ev Maoz and Amnon Rapoport, "The Condorcet-Efficiency of Sophisticated Voting Under the Plurality and Approval Procedures", University of Haifa, Institute of Information Processing and Decision Making, Report No. 76, January 1990.
12. Amnon Rapoport, Dan S. Felsenthal and Ze'ev Maoz, "Sincere Versus Strategic Behavior in Small Groups", University of Haifa, Institute of Information Processing and Decision Making, Report No. 80, April 1990.

13. Dan S. Felsenthal, Ze'ev Maoz and Amnon Rapoport, "Does It Really Make a Difference? An Empirical Evaluation of Six Voting Procedures", University of Haifa, Institute of Information Processing and Decision Making, Report No. 103, May 1992.
14. Book Review on Abraham Diskin's book *Elections and Voters in Israel*, (New York: Praeger, 1991); Reviewed in: *Studies in Contemporary Jewry* 8 (June 1992), 374–376.
15. Vincent C.H. Chua and Dan S. Felsenthal, "Coalition Formation Theories Revisited: An Empirical Investigation of Aumann's Hypothesis", London School of Economics and Political Science, Centre for the Philosophy of the Natural and Social Sciences, Voting Power Report VPP 06/xx, revised August 2006. Downloadable from: http://eprints.lse.ac.uk/767/
16. Dan S. Felsenthal, "Maintaining OPOV: Some Problems Associated with the Periodic Need to Re-Draw Parliamentary Boundaries in the UK and the US", London School of Economics and Political Science, Centre for the Philosophy of the Natural and Social Sciences, Voting Power Report 08/01, May 2008. Downloadable from: http://www.lse.ac.uk/collections/VPP/VPPpdf/VPPPublications/VPP08_04.pdf
17. Dan S. Felsenthal, "Review of Paradoxes Afflicting Various Voting Procedures Where One Out of m (m ^{3} 2) Candidates Must Be Elected", London School of Economics and Political Science, Centre for Philosophy of Natural and Social Science, Symposium on *Assessing Alternative Voting Procedures*, 27 May 2010. Downloadable from http://eprints.lse.ac.uk/27685/
18. Dan S. Felsenthal and Nicolaus Tideman. "Weak Condorcet Winner(s) Revisited", February 2013.
19. Dan S. Felsenthal and Nicolaus Tideman. "Interacting Double Monotonicity Failure with Direction of Impact under Five Voting Methods", March 2013.

Moshé Machover: List of Publications

[Note: This list excludes joint publications with Dan Felsenthal]

A. Books

1. John L Bell and Moshé Machover, *A Course in Mathematical Logic*, North-Holland, 599 pp, 1977, Second printing 1986.
2. Emmanuel Farjoun and Moshé Machover, *Laws of Chaos*: *A Probabilistic Approach to Political Economy*, Verso, 264 pp, 1983.
3. Moshé Machover, *Set Theory, Logic and their Limitations*, Cambridge University Press, ix + 288 pp, 1996, Second printing 1998.

B. Papers in Refereed Journals

1. Moshé Machover, "Note on sentences preserved under direct products and powers". *Bull de l'Acad Polonaise des Sciences, Ser Sc Math*, **3** (1960) 519–23.
2. —, "The theory of transfinite recursion". *Bull American Math Soc*, **67** (1961) 575–8.
3. —, "Contextual determinacy in Lesniewski's grammar". *Studia Logica*, **19** (1966) 47–57.
4. I Juhász and Moshé Machover, "A note on non-standard topology", *Indag Math*, **31** (1970) 482–4.
5. Moshé Machover, "Towards a new philosophy of mathematics". *British Journal for the of Philosophy of Science*, **34** (1983) 1–11.
6. E Farjoun and Moshé Machover, "Probability, economics and the labour theory of value", *New Left Review* **152** (1985) 95–108.
7. Moshé Machover, "The place of nonstandard analysis in mathematics and in mathematics teaching", *British Journal for the Philosophy of Science*, **44** (1993) 205–12.
8. G Garret, I MacLean and Moshé Machover, "Power, power indices, and blocking power: A comment on Johnston", *British Journal of Political Science*, **25** (1995) 563–568.

R. Fara et al. (eds.), *Voting Power and Procedures*, Studies in Choice and Welfare,
DOI 10.1007/978-3-319-05158-1_23,

9. Moshé Machover, "Notions of a priori voting power: Critique of Holler and Widgrén", *Homo Oeconomicus*, **XVI(4)** (2000) 415–25. Also *Notizie di Politea*, **59** (2000) 29–38.
10. —, "Tony Blair's dilemma", *Homo Oeconomicus* **XIX(4)** (2003) 576–581.
11. I Lindner and Moshé Machover, "L S Penrose's limit theorem: Proof of some special cases", *Mathematical Social Sciences* **47** (2004) 37–49.
12. M Hosli and Moshé Machover, "The Nice Treaty and voting rules in the Council: A reply to Moberg", *Journal of Common Market Studies*, **42** (2004) 497–521.
13. P-L Chang, V C H Chua and Moshé Machover, "L S Penrose's limit theorem: Tests by simulation", *Mathematical Social Sciences* **51** (2006) 90–106.
14. —, "The underlying assumptions of electoral systems" in Felsenthal D S and Machover M (eds.) *Electoral Systems: Paradoxes, Assumptions, and Procedures*, Springer, (2012) 3–9.

C. Articles in Books (Other Than Conference Proceedings)

1. Moshé Machover, "Abraham Robinson" (Introduction to the correspondence between Kurt Gödel and A Robinson) in: Solomon Fefferman et al. (eds.) *Kurt Gödel Collected Works* Volume V, Oxford University Press, (2003) 191–194.
2. —, "The stochastic concept of economic equilibrium: a radical alternative" In: D DeVidi, M Hallett and P Clarke, (eds.) *Logic, Mathematics*, *Philosophy, Vintage Enthusiasms: Essays in Honour of John L. Bell*. The Western Ontario series in philosophy of science (75). Springer, Dordrecht (2011) 413–421. ISBN 9789400702134 http://eprints.lse.ac.uk/36428

D. Articles in Conference Proceedings

1. Moshé Machover, "Comment on Schmidtchen and Steunenberg", in: Holler M. J., Kliemt H, Schmidtchen D and Streit M. E., (eds.) *Power and Fairness* (Jahrbuch für Neue Politische Ökonomie Vol 20); Mohr Siebeck, (2002) 225–227.
2. —, "Comment on Braham and Steffen", in: Holler M.J., Kliemt H, Schmidtchen D and Streit M. E., (eds.) *Power and Fairness* (Jahrbuch für Neue Politische Ökonomie Vol 20); Mohr Siebeck, (2002) 349–353.
3. D Leech and Moshé Machover, "Qualified majority voting: The effect of the quota", in: Holler M.J., Kliemt H, Schmidtchen D and Streit M. E., (eds.) *European Governance* (Jahrbuch für Neue Politische Ökonomie Vol 22); Mohr Siebeck, (2003)127–143.
4. Moshé Machover, "A forgotten common origin: Comments on Hegselmann", in: Gillies, D, (ed) *Laws and Models in Science*; King's College Publications, Mohr Siebeck, (2005) 47–49.
5. —, "Penrose's square-root rule and the EU Council of Ministers: significance of the quota". In: Cihocki M. A. and Życzkowski K. (eds.) *Institutional Design and Voting Power in the European Union*; Ashgate, (2010) 35–42. http://eprints.lse.ac.uk/2857/

E. Encyclopedia Entry

1. Moshé Machover, "Analysis, Non-Standard". 3000-word entry in the *Routledge Encyclopedia of Philosophy* (1997)

F. Reports and Book Reviews

1. Moshé Machover, *Theory of Transfinite Recursion*. (Hebrew) PhD Thesis, Hebrew University of Jerusalem, 1962.
2. —, *Decision Procedure for Wellformedness in Lesniewski's Grammar*. Published as Report #16 of Applied Logic Branch, HUJ, Israel, 1964, 15 pp.
3. —, *Non-standard Analysis without Tears*. Published as Report #27 of Applied Logic Branch, HUJ, Israel, (1967) 56 pp.
4. —, *Projective Geometry*. Israel Centre for Science Education, Jerusalem, Israel, (1968) 47 pp.
5. —, *Axiomatic Arithmetic*. Israel Centre for Science Education, Jerusalem, Israel, (1968) 30 pp.
6. — and J Hircshfeld, *Lectures on Non-Standard Analysis*, #94 in the series Lecture Notes in Mathematics, Springer, (1969) 79 pp.
7. —, Review of *Recursive Aspects of Descriptive Set Theory* by R. Mansfield and G. Weitkamp. *Bull LMS* **18** (1986) 429–430.
8. —, Review of *Varieties of Constructive Mathematics* by D. Bridges and F. Richman. *Bull LMS* (1988).
9. —, Review article on *The Bounds of Logic. A Generalized Viewpoint* by G. Sher. *British Journal Philosophy of Science* **45** (1994) 1078–1083.
10. —, Review of *Abraham Robinson: The Creation of Nonstandard Analysis — A Personal and Mathematical Odyssey* by Joseph W. Dauben. *British Journal Philosophy of Science* **47** (1996) 137–140.
11. —, Review of *Real Numbers, Generalizations of Reals, and Theories of Continua*, Philip Ehrlich (ed.). *British Journal for the Philosophy of Science* **47** (1996) 320–324.
12. —, Interview (19.11.97) for "Decision-making in the EU Council of Ministers", a radio documentary, broadcast 5.12.97 on the Radio of the Finnish Broadcasting Company.
13. —, "A Nice Trap" (Analysis of the articles in the Treaty of Nice dealing with the decision rule of the EU Council of Ministers). Published in *Polarities* (Newsletter of the Finnish Institute), 2001.
14. —, Review of *Simple Games, Desirability Relations, Trading, Pseudoweightings* by A Taylor and W Zwicker. *Social Choice and Welfare* **18** (2001) 617–618.
15. —, "Discussion topic: voting power when voters' independence is not assumed", presented at workshop Voting Power in Practice, 29–31 August 2007, Warwick University, UK. http://eprints.lse.ac.uk/2966/
16. —, *Collective Decision-Making and Supervision in a Communist Society*, (2009) 49 pp. http://eprints.lse.ac.uk/51148/
17. — and S. Terrington, "Mathematical structures of simple voting games" (2013) http://eprints.lse.ac.uk/47850/