Electoral Systems

Paradoxes, Assumptions, and Procedures

Springer

Contents

Part I
Representative Electoral Systems: Underlying Assumptions and Decision Rules

Chapter 1
The Underlying Assumptions of Electoral Systems

Moshé Machover

1.1 Introduction

My aim in this brief paper is modest: not to present new findings, but to propose what I regard as a useful way of classifying voting procedures, and thus organizing the way we look at them. My main thesis is that we have to make a strict distinction between two kinds of consideration in choosing a voting/election procedure:

- *Political criteria*. I use this rubric in a very broad sense, including criteria ranging from the pragmatic to the philosophical. But all of them are purely a matter of opinion, not of "right" or "wrong".
- *Social-choice considerations*. I take this rubric in the narrow sense: the logico-mathematical properties of a voting procedure, the pathologies and paradoxes that afflict it.

These two kinds of consideration are not on a par with each other: political considerations are paramount in choosing a voting procedure. For example, as far as political elections are concerned, it is politicians who usually choose the voting procedure; and even when the choice is made by referendum, the question put to referendum is framed by politicians. But politicians and their advisors – and, ideally, the general public – ought to be aware of the logico-mathematical properties of the voting procedures in question; otherwise they can easily walk into a trap. So it is wrong to dismiss these matters as of interest only to geeks.

On the other hand, social-choice theorists must recognize that their professional scientific role is confined to ascertaining the technical properties of voting procedures, including the likelihood of various pathologies manifesting themselves under each procedure. However, the decision as to which pathology (with a given

M. Machover (✉)
Centre for the Philosophy of the Natural and Social Sciences, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK
e-mail: moshe.machover@kcl.ac.uk

D.S. Felsenthal and M. Machover, *Electoral Systems*, Studies in Choice and Welfare,
DOI 10.1007/978-3-642-20441-8_1,

likelihood) is more tolerable than another is not a scientific matter; it is political. On this, the opinion of a social-choice theorist is not more privileged than that of any well-informed member of the public.

Similarly, social-choice theory provides information about the effects of a given system of electing a legislature regarding the stability of a parliamentary government and the number of parties with realistic prospect of winning seats (Duverger's law). However, the question as to the importance of a stable government (and the desirable *degree* of stability), or the desirability of a small or large number of such parties is a purely political one.

In what follows, I will formulate two main political dichotomies, each offering two alternatives. This gives rise to a fourfold political classification of voting procedures. I will explore what social choice theory has to offer in each of these four classes.

1.2 Two Dichotomies

The first main dichotomy is relevant for electing a representative assembly such as a legislature, not a holder of an individual post, such as a president. I state it as follows:

i. Proportional Representation (PR) vs. District Representation (DR)

This dichotomy hinges on a distinction between two quite different senses of the verb *represent* and its derivatives. Who or what is being "representative", and whom or what are they supposed to "represent"?

One sense of this term – which underlies PR – is intended, for example, in statistics, when we speak of a *representative sample*. An elected assembly is representative in this sense if it is a microcosm of the entire electorate, reflecting in true proportion (or as near to it as possible) the various shades of opinion that exist in the society as a whole. Thus it can stand as proxy for a market-place meeting of the entire citizenry; and a vote taken in the assembly may be regarded as a close approximation to a referendum. Here a member of the assembly does not represent a geographically defined constituency, but reflects a like-minded section of the electorate at large, which may well be geographically dispersed. Note that being representative in this sense is primarily an attribute of the assembly as a body, not so much of each individual member: in order to ascertain whether the assembly is indeed representative, we must examine it as a whole.

Another, quite different sense of the term – which underlies DR – is similar to the one intended when we speak of a *diplomatic representative* of a country. Note that being *a representative* in this sense is an attribute of the individual member: an assembly is representative only inasmuch as it is an assembly *of* representatives. The relationship between a representative in this sense and what s/he represents is like that between agent and principal.[1] Here every member of the assembly is personally elected for representing a particular constituency, which is usually defined geographically. Accordingly, there are a large number of constituencies, each of which elects a single representative or a small number – at most a handful – of representatives. Naturally, such a constituency may be, and normally is in fact, quite heterogeneous: its voters may differ considerably from one another in their interests, preferences, tastes and opinions. The presumed aim of a DR procedure is to elect a candidate (or a small set of candidates) that is in some sense "best" or "most suitable" for representing this heterogeneous constituency.

Although the distinction between these two senses of representation is quite fundamental, I have not seen it clearly and explicitly articulated in the social-choice literature. Perhaps this is due to my ignorance; and I stand to be corrected. At any rate, the distinction is very often ignored and the two senses are conflated.[2]

However, there are some well-known "compromise" systems that blend both types of representation. One such compromise is the so-called Additional Member system used, for example, in elections to the German Bundestag and the Scottish Assembly, whereby some members of the legislature are elected by a DR method, and the rest are elected by a PR method, designed to achieve or approach overall proportionality. A second, quite different compromise consists in dividing the electorate at large into fairly large geographically-based constituencies, within each of which elections are held using PR. This compromise is used, for example, in the UK in elections to the European Parliament; it has occasionally been used in elections to the French National Assembly.

The second main dichotomy is:

ii. Deterministic Processing (DP) vs. Lottery Processing (LP)

Here "processing" refers to the way the votes cast are processed to produce the outcome of the election.

I consider a voting procedure to be DP even if it does use lottery, provided this use is confined to resolving ties, whose occurrence is extremely unlikely. Thus an LP procedure is one that relies on lottery in a major way.

Whether use of LP is acceptable is clearly a political matter (in the broad sense) and depends on social norms and on the purpose for which the election is conducted. According to current social norms, it is considered in many countries desirable to select a trial jury by lot out of a large pool of admissible candidates. But electing a legislature by LP would probably be regarded by most people as unacceptable. Electing an individual by lottery for a position such as chairman of a meeting is quite common, but electing a holder of high political office by LP would be unacceptable – although it was normal practice under Athenian democracy.

---

[1] I owe this observation to Iain McLean (oral communication).

[2] This goes back to John Stuart Mill. In (Mill 1861, Chap. 7) he clearly advocates PR; but then seems to take it for granted that electing a legislature must use some form of DR.

1.3 PR Procedures

Let us now see what social choice has to offer if we opt for PR.

1.3.1 PR and DP

The only electoral procedure that really implements this combination (as far as possible) is the list system. To be precise, there are two variants of this system. In the *closed* list variant, the seats are allocated to a party's candidates in the order in which they appear on its list. In the *open* list variant, voters may indicate preference for a particular candidate in the list of their choice, and seats are allocated accordingly.[3]

The STV procedure is often claimed by politicians and journalists to be a PR system. But social-choice theorists know very well that this claim is incorrect. This is not only easy to prove in theory (for example, by observing that STV is not monotonic), but can also be seen in practice by examining the results of elections conducted under STV.[4] In fact, STV is a DR system that is ingeniously designed to produce less disproportionate outcomes than the extremely pathological plurality procedure.[5] However, the approximate degree of proportionality it produces is quite erratic. In particular, STV is biased against small and radical parties.

1.3.2 PR and LP

There is one – and as far as I know only one – procedure that implements this combination of political alternatives. It is the lottery voting procedure (LVP) proposed by the American jurist and political scientist Akhil Reed Amar (1984).[6] This is how it works. The entire electorate is divided into constituencies of roughly equal size. Elections are conducted in each constituency as under the plurality system, but with the following crucial difference. Whereas under the plurality system the winner is the candidate with the greatest number of votes, under LVP a weighted lottery is conducted, with candidates' weights proportional to the respective numbers of votes cast for them.

---

[3] Note however that the aggregation of all the individual preference orderings into a single overall ordering is problematic, due to Arrow's Theorem.

[4] For example, see http://en.wikipedia.org/wiki/Dail for results of elections to the Irish Dáil.

[5] STV is therefore advocated by people who can see the virtues of PR, but are wedded to DR either on political grounds or because they simply take it for granted. Among the latter was J S Mill (1861, Chap. 7); cf. footnote 2.

[6] It is also known, somewhat misleadingly, as the "random dictator" procedure.

Using Kolmogorov's Strong Law of Large Numbers, it is not difficult to show that the overall outcome under LVP is almost certain to be extremely close to proportionality. More precisely, if the number of constituencies is fairly large (say 100 or more) then the total number of seats won by candidates representing a given party or informal trend of opinion is very highly likely to be closely proportional to the total number of votes cast at large for such candidates.[7]

This procedure shares some of the attractive political properties of both deterministic PR and DR.[8] In fact, superficially, LVP looks like a DR procedure, but this is not really so. The winner of the election in a given constituency is not supposed to be its "best" or "most suitable" representative. In fact, her or his primary allegiance is not to the constituency but to the party or trend of opinion for which s/he stands. The constituency serves primarily as a subspace of the sampling space of the electorate at large. Indeed, in principle there is no need for the constituencies to be determined geographically; they can be quite arbitrary sections, roughly equal in size, of the electorate at large. (However, this would destroy some important political advantages of LVP.)

1.4 DR Procedures

Here things will get somewhat messy. But before that, I would like to introduce a subsidiary dichotomy, singling out a particular political principle:

iii. *Majority Rule (MR) vs. Aggregation Rules (AR)*

MR systems are based on the political view that regards majority rule as a paramount principle. The meaning of MR is clear enough when there are just two candidates. The straightforward natural generalization of this is *Condorcet's Principle*:

> If candidate x dominates candidate y (i.e., x is preferred to y by a majority of the voters), then x is socially preferable to y.

Note that in order to apply this rule, it is not necessary in principle for a voter to order the candidates in a (transitive) preference ordering. Only pairwise comparisons are needed. And a voter's comparisons may contain cycles. (It is sometimes claimed that cyclic preferences are irrational. I don't find this claim persuasive. Besides, is it politically acceptable to disqualify or ignore voters whose voting behaviour is allegedly irrational? That would be extremely dangerous. . . .)

The alternative to MR is a mixed bag of various rules for aggregating degrees of approval (or preference) that are assigned by the voters to each candidate. These "degrees" may be ordinal, cardinal or of an intermediate kind (as in grading by

[7] For a proof, see Machover (2009, Sect. 6.2).

[8] For a discussion of the technical properties and political advantages of this procedure, see Amar (1984) and Machover (2009, Sect. 4.4).

marks that are not merely ordinal, but are not reducible to cardinal numbers). But in any case they require or imply at least a transitive weak ordering of the candidates by each voter.[9]

Aggregation systems pose two distinct problems. First, can degrees of approval (or preference) assigned by different voters be meaningfully aggregated? This problem is familiar in relation to utilities; but it is more general.

Second, aggregation involves loss of information: in general, the voting profile contains much more information than the outcome of the election. Arrow's theorem is a particular manifestation of this: it applies only to procedures that try to aggregate ordinal preferences (preference orderings) into a single "social" ordering. However, the problem is more general.

This loss of information can be regarded as the source of all the paradoxes and pathologies that afflict voting procedures. I will not discuss these matters any further, but refer you to Dan Felsenthal's paper (Chap. 3 of this volume).

Let me just add that as far as I know the problems posed by the paradoxes and pathologies of AR procedures arise whether we insist on deterministic processing (that is, the combination AR and DP) or allow lottery processing (that is, AR and LP).

The situation regarding MR is different – which is the reason I have singled it out in the subsidiary dichotomy (iii).

The combination MR and DP needs to be supplemented by some method of aggregating preferences, in case a Condorcet winner does not exist. Thus we are back to the problems raised in the case of the combination AR and DP.

This leaves the final combination:

1.4.1 MR and LP

For this combination, if just one candidate needs to be elected, social-choice theory provides an elegant unique optimal solution, and does not need to be supplemented by any other political principle.

This solution is provided by a beautiful theorem, proved in 1991 by Laffond et al. (1993), and independently (using a quite different method) by Fisher and Ryan (1992).

Let me outline this theorem. Consider the following *tournament game*: a two-person game in which each of two players, I and U, must nominate (independently of each other) one member of the set X of candidates standing for election. Suppose I nominates x and U nominates y. If $x \succ y$ (i.e., if x dominates y), then U pays I \$1; if $y \succ x$, then I pays U \$1; and if $x = y$ no payment is made.

[9] A very rudimentary marking is used in the plurality and approval voting procedures, where the only admissible marks are 0 and 1.

The theorem states that in this game there is a unique optimal mixed strategy. In other words, there are unique probabilities $\{p_x : x \in X\}$, with $\sum_{x\in X} p_x = 1$, such that if a player uses a lottery with these probabilities to nominate a candidate, then s/he maximizes her/his expected payoff. (By symmetry, this maximal payoff is of course 0.) Clearly, the support of this probability distribution (the set $\{x \in X : p_x > 0\}$) is a subset of the top cycle of candidates. In particular, if there is a Condorcet winner, the optimal strategy is pure, and assigns that candidate probability 1. Rather surprisingly, the support always consists of an odd number of candidates.

As pointed out by Felsenthal and Machover (1992), this provides an electoral procedure based purely on MR and LP: conduct a weighted lottery, in which each candidate x is assigned weight p_x.

1.5 Conclusion

Much of social-choice literature is concerned with the perplexing problematics of selecting an acceptable election procedure out of a large number of competing ones. What I have tried to show is that if one subscribes to certain simple "grand" *political* options, or a combination of these, then social choice can provide a single optimal procedure.

Acknowledgements Valuable comments from Dan S Felsenthal and Maurice Salles are gratefully acknowledged.

References

Amar, A. R. (1984). Choosing representatives by lottery voting. *Yale Law Journal, 93*, 1283–1308.

Felsenthal, D. S. & Machover, M. (1992). After two centuries, should Condorcet's voting procedure be implemented? *Behavioral Science, 37*, 250–274.

Fisher, D. C. & Ryan, J. (1992). Optimal strategies for a generalized 'Scissors, Paper and Stone' game. *American Mathematical Monthly, 99*, 935–942.

Laffond, G., Laslier, J., & Le Breton, M. (1993). The bipartisan set of a tournament. *Games and Economic Behavior, 5*, 182–201.

Machover, M. (2009). *Collective decision-making and supervision in a communist society*. Mimeograph, downloadable from http://tinyurl.com/ydfbs5z.

Mill, J. S. (1861). *Considerations on representative government*. Downloadable from http://www.gutenberg.org/ebooks/5669

Chapter 2
Some Informal Remarks on Devising a "Fair" Decision-Making Rule for Representative Assemblies

Dan S. Felsenthal

In the previous chapter Moshé Machover distinguished between two kinds of representative assemblies, each of which can be elected by using either a deterministic or a probabilistic voting procedure:

1. A *PR assembly* which is a microcosm of the entire electorate and where every member represents an ideologically homogeneous but geographically dispersed constituency. Machover argues – and I agree – that the only way to obtain such an assembly by using a deterministic voting procedure is to use the closed list system procedure.
2. A *DR assembly* where every member represents (is an agent of) an ideologically diverse but geographically contiguous constituency, of which s/he is in some sense the "best" representative.

However, Machover did not address an important related issue, i.e., how to devise a "fair" decision-making rule to be used by the assembly. I would like to dwell on this issue.

As far as I know, it is a universal practice in democracies that the decision rule used by both types of assembly, as well as the decision rule used in popular referenda, is majoritarian.[1] The fact that democracies employ decision-making rules within representative assemblies that are based solely on the majority principle could lead to the conclusion that a decision based on this principle can always

[1] A majoritarian decision rule requires that in order to change the status quo slightly more than half the voters and fewer than all the voters must support this change. Note that since a requirement of unanimity enables every voter to veto a change of the status quo, unanimity is not considered a majoritarian voting rule.

D.S. Felsenthal (✉)
School of Political Sciences, University of Haifa, Israel

Centre for Philosophy of Natural and Social Science, London School of Economics, UK
e-mail: dfelsenthal@poli.haifa.ac.il

D.S. Felsenthal and M. Machover, *Electoral Systems*, Studies in Choice and Welfare,
DOI 10.1007/978-3-642-20441-8_2,

be reconciled with democratic principles. In fact, "there is nothing inherent in democracy that requires majority rule." (Guinier 1994, p. 17). Moreover, the majority rule principle implies that the minority of representatives in an assembly are unable to affect reality – even if their number is exactly proportional to the proportion of the electorate who supported them in the election.

So the universal use of the majoritarian decision rule within representative assemblies, as well as in popular referenda, causes the minority to become totally impotent in shaping public policies. Such impotence is especially serious if the minority group is relatively large and it is a permanent one, i.e., it always consists of the same (type of) voters who may belong to the same ethnic group, or the same ideological party, or the same geographic region.

In contemplating which decision rule(s) ought to be used by a representative assembly when it engages in tasks involving the selection of one out of several possible alternatives, let us first consider three alternative political/philosophical principles or goals:

1. *Majority rule*: To guarantee that the alternative preferred by the majority of voters (or representatives) will be selected.
2. *Equiprobability of success*: To let every voter (or representative) have the same probability that his/her most preferred alternative is the one selected.
3. *Equal opportunity to avoid the worst*: To provide every voter (or representative, or alliance) with an ability to prevent that his/her/its least preferred alternative is the one selected.

To achieve the first goal one must use a deterministic voting procedure and select the alternative supported by the majority of voters or representatives.

To achieve the second goal one must use a probabilistic voting procedure which assigns to every voter or representative the same chance of being selected – and the selected voter/representative will then state which alternative s/he prefers.[2]

To achieve the third goal one must enable every voter, or group of voters of some minimal size, to veto one of the alternatives under consideration, thus guaranteeing that every voter or group of voters has some minimal effect on the selected outcome.

Achieving the first goal can never make, by definition, the majority of voters very miserable, but it may make the minority of voters very miserable. On the other hand, achieving the second goal may make either the majority or the minority very miserable if a member of the other group is selected to choose the alternative to be implemented. (Of course there is a higher probability that the selected person will belong to the majority than to the minority group.)

It seems natural that we prefer the possibility that the minority may be miserable over the possibility that the majority may be miserable – and hence we prefer to realize the first goal (principle) over the realization of the second goal. Moreover, since the realization of the second goal involves the employment of a probabilistic

[2] In an assembly made up of party-blocs, the achievement of this goal would require the probability of success of each bloc to be proportional to its weight.

voting procedure, this procedure may cause additional problems such as instability and/or inconsistency in public policies and decision-making.[3]

The realization of the third goal has two advantages over the realization of the other two goals: first, it leads to the selection of an alternative that is stable regardless of whether the social preference ordering among the available alternatives contains cycles.[4] Second, if voters behave rationally, then the selected alternative is not only Pareto-optimal – that is, no other alternative is preferred by all the voters over some other alternative – but it also does not constitute any voter's least-preferred alternative.

It is quite easy to realize the third goal when the number of voters/representatives (n) is smaller than the number of policy alternatives (m) of which one or more must be selected. In this case every voter/representative, in his/her turn, can veto one or more alternatives (depending on his/her weight) – and the alternative(s) that was/were not vetoed is/are selected.[5] However, implementing this goal is difficult when the number of representatives in the assembly is larger than the number of policy alternatives. Implementing this goal in this case – which is common in actual representative assemblies – implies that more than one representative is needed to veto any given policy alternative, and the formation of the needed alliance(s) may become complicated. Moreover, a satisfactory theory as to how to analyze such cases is still lacking.

So how, if at all, is it possible to adopt one or more decision rules in representative assemblies which will provide the representatives belonging to the minority group with both a priori as well as actual voting power proportional to their weight?

I think the answer to this question is twofold:

(a) Act according to a fourth political-philosophical principle or goal, i.e.,

4. *Proportionality of a priori voting power to weight*: Let every voter or representative have the same probability of being critical in a division. Given that the number of seats controlled by the various parties or geographical units in a representative assembly is proportional to the number of relevant

[3] Instability may be caused by the losers' demand following any given division to conduct another round of voting, whereas inconsistency may be a result of adopting contradictory policies that were selected interchangeably by voters belonging to the majority and minority groups.

[4] A stable alternative is an alternative that cannot profitably be objected to by any voter or alliance of voters. Note that, in contrast, no alternative is stable if one uses a majoritarian decision rule and the social preference ordering among the available alternatives contains a top cycle.

[5] This procedure is known as sequential voting by veto (SVV). It was proposed originally by Mueller (1978) who presented an algorithm for determining the winning alternative under SVV, given the order in which the voters/representatives cast their vetoes. Moulin (1981, 1983, pp. 138–140) extended Mueller's idea to any situation in which n voters have to select one out of $n + 1$ alternatives and they have complete information on all other voters' preference orderings among the alternatives. Felsenthal and Machover (1992) generalized the Mueller–Moulin result to a situation in which n voters/representatives must select s out of m alternatives ($s > 0; m > n \geq 2$). For laboratory experiments with small groups operating under SVV see Yuval (2002) and Yuval and Herne (2005).

> voters in the electorate,[6] one looks for a decision rule to be used by the assembly such that the *a priori relative voting power* (as measured by Banzhaf's index of relative voting power) will be as close as possible to each representative's relative weight, i.e., the proportion of voters s/he represents.[7]

However, the realization of this goal too is not problem-free. To understand just one of the problems associated with it, consider the following simple example.

Suppose a 99-seat legislative assembly with three parties, each controlling 33 seats because each received an equal proportion (1/3) of the votes in an election. The *relative* a priori voting power of each party will be 1/3 – which is proportional to each party's weight – regardless of whether the quota (q) needed to pass resolutions is $34 \leq q \leq 66$ (simple or qualified majority) or $67 \leq q \leq 99$ (unanimity). However, the *absolute* a priori voting power of every party (as measured by the Penrose measure) would be 1/2 if only simple or qualified majority would be required to pass a resolution, but only 1/4 if unanimity is needed to pass a resolution. So which quota would be fairer in this case? It would seem that here unanimity would be fairer than simple/qualified majority, because under simple/qualified majority any single party may become powerless if the other two parties form a relatively long-term binding alliance. Of course the price to be paid for granting in this case veto power to each of the three parties, is not only significant loss of (a priori) absolute voting power by each of the parties, but also the possibility of total paralysis of the legislature inasmuch as the parties are unable to agree on the passage of any bill.

(b) So perhaps some milder form of de facto (proportional) voting power would be preferable than awarding each of the three parties in the above example veto power regarding all proposed bills. It should be possible to institute arrangements, at least with respect to certain kinds of decisions, e.g., budgetary decisions, or decisions regarding certain regions or policy areas, which will enhance the a posteriori (actual) voting power of representatives belonging to the minorities in legislatures. For example, if the Red and Blue parties control 40% and 60% of the seats, respectively, in a representative assembly, then one can institute an arrangement where the Blue party would be given the prerogative of determining the total size of the annual budget, as well as dividing it into parts – one containing 60% of the total planned expenditure and the other containing 40% of the total expenditure – and let the Red party have the sole prerogative to decide how the 40% part of the budget would be allocated.

Of course such power-sharing arrangements – in decision-making bodies in general and legislatures in particular – are not problem-free, and their details are crucial. Yet it seems to me that the discussion and development of such proportional

[6] The relevant electorate is either the number of voters who supported each party represented in the assembly, or the number of voters belonging to each geographical unit represented in the assembly, or the size of financial contribution of each member-country to the common fund, e.g., the financial contribution of each member of the International Monetary Fund to its fund.

[7] For the various measurements of a priori voting power see Felsenthal and Machover (1998).

power-sharing arrangements should constitute a major new area for social choice theory to be engaged in.

So far the relevant disciplines (social choice, political science, economics, mathematics, law, philosophy) focused mainly on how to elect a representative assembly, and to a lesser degree on how to measure the a priori voting power of representatives in an assembly. In my opinion the time has come for them to shift their focus to how to devise a fair and practical decision rule(s) for a representative assembly so that *all* its members will have *actual* voting power which is as close as possible to their relative weight.

Acknowledgements I wish to thank Moshé Machover and Maurice Salles for their useful comments on these remarks.

References

Felsenthal, D.S., & Machover, M. (1992). Sequential voting by veto: extending the applicability of the Mueller–Moulin algorithm. *Theory and Decision, 33*, 223–240.

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

Guinier, L. (1994). *The tyranny of the majority: fundamental fairness in representative democracy*. New York: Free Press.

Moulin, H. (1981). Prudence versus sophistication in voting strategy. *Journal of Economic Theory, 24*, 398–412.

Moulin, H. (1983). *The strategy of social choice. North Holland: Amsterdam*.

Mueller, D.C. (1978). Voting by veto. *Journal of Public Economics, 10*, 55–75.

Yuval, F. (2002). Sophisticated voting under the sequential voting by veto. *Theory and Decision, 53*, 343–369.

Yuval, F., & Herne, K. (2005). Sophisticated behavior under majoritarian and non-majoritarian voting procedures. *Political Behavior, 27*, 217–237.

Part II
Paradoxes Afflicting Electoral Procedures and Their Expected Probability

Chapter 3
Review of Paradoxes Afflicting Procedures for Electing a Single Candidate

Dan S. Felsenthal

3.1 Introduction

Three factors motivated me to write this chapter:

- The recent passage (25 February 2010) by the British House of Commons of the *Constitutional Reform and Governance Bill*, clause #29 of which states that a referendum will be held by 31 October 2011 on changing the current *single member plurality* (aka *first-past-the-post*, briefly FPTP) electoral procedure for electing the British House of Commons to the (highly paradoxical) *alternative vote* (AV) procedure (aka *Instant Runoff*).[1] Similar calls for adopting the alternative vote procedure are voiced also in the US.
- My assessment that both the UK and the US will continue to elect their legislatures from single-member constituencies, but that there exist, from the point of view of social-choice theory, considerably more desirable voting procedures for electing a single candidate than the FPTP and AV procedures.
- A recent report by Hix et al. (2010) – commissioned by the British Academy and entitled *Choosing an Electoral System* – that makes no mention of standard social-choice criteria for assessing electoral procedures designed to elect one out of two or more candidates.

[1]Following the general elections held in the UK on 6 May 2010, a coalition government has been formed between the Conservative and Liberal-Democratic parties in which the two parties committed to hold a referendum on the possible change of the election procedure to the House of Commons from FPTP to AV. In the referendum held on 5 May 2011 it was decided to keep the FPTP procedure.

D.S. Felsenthal (✉)
School of Political Sciences, University of Haifa, Israel

Centre for Philosophy of Natural and Social Science, London School of Economics, UK
e-mail: dfelsenthal@poli.haifa.ac.il

D.S. Felsenthal and M. Machover, *Electoral Systems*, Studies in Choice and Welfare,
DOI 10.1007/978-3-642-20441-8_3,

I therefore thought it would be well to supplement that report by reminding social choice theorists, political scientists, as well as commentators, policymakers and interested laymen – especially in the UK and the US – of the main social-choice properties by which voting procedures for the election of one out of two or more candidates ought to be assessed, and to list and exemplify the paradoxes afflicting these voting procedures.

Thus this paper should be regarded as an updated review by which to assess from a social-choice perspective the main properties of various known voting procedures for the election of a single candidate.

Of the 18 (deterministic) voting procedures analyzed in this paper, the Condorcet-consistent procedures proposed by Copeland (1951) and by Kemeny (1969) seem to me to be the most desirable from a social-choice perspective for electing one out of several candidates.

The paper is organized as follows: In Sect. 3.2 I survey 15 paradoxes, several of which may afflict any of the 18 voting procedures that are described in Sect. 3.3. Section 3.4 summarizes and presents additional *technical-administrative* criteria which should be used in assessing the relative desirability of a voting procedure. In the detailed appendix in Sect. 3.5 I exemplify most of the paradoxes to which each of the surveyed election procedures is susceptible.

3.2 Voting Paradoxes

I define a "*voting paradox*" as an undesirable outcome that a voting procedure may produce and which may be regarded at first glance, at least by some people, as surprising or as counter-intuitive.

I distinguish between two types of voting paradoxes associated with a given voting procedure:

1. "*Simple*" or "*Straightforward*" paradoxes: These are paradoxes where the relevant data leads to a "surprising" and arguably undesirable outcome. (The relevant data include, *inter alia*, the number of voters, the number of candidates, the number of candidates that must be elected, the preference ordering of every voter among the competing candidates, the amount of information voters have regarding all other voters' preference orderings, the order in which voters cast their votes if it is not simultaneous, the order in which candidates are voted upon if candidates are not voted upon simultaneously, whether voting is open or secret, the manner in which ties are to be broken).
2. "*Conditional" paradoxes*: These are paradoxes where changing one relevant datum while holding constant all other relevant data leads to a "surprising" and arguably undesirable outcome.

An array of paradoxes of one or both types are described and analyzed by McGarvey (1953), Riker (1958), Smith (1973), Fishburn (1974, 1977, 1981, 1982), Young (1974), Niemi and Riker (1976), Doron and Kronick (1977), Doron (1979), Richelson (1979), Gehrlein (1983), Fishburn and Brams (1983), Saari (1984, 1987, 1989, 1994, 2000, 2008), Niou (1987), Moulin (1988a), Merlin and Saari (1997), Brams, Kilgour and Zwicker (1998), Scarsini (1998), Nurmi (1998a, 1998b, 1999, 2004, 2007), Lepelley and Merlin (2001), Merlin et al. (2002), Merlin and Valognes (2004), Tideman (1987, 2006), Gehrlein and Lepelley (2011), among others.

3.2.1 Simple Paradoxes

The six best-known "simple" paradoxes that may afflict voting procedures designed to elect one out of two or more candidates are the following:

3.2.1.1 The Condorcet (or Voting, or Cyclical Majorities) Paradox (Condorcet 1785; Black 1958)

Given that the preference ordering of every voter among the competing candidates is transitive, the (amalgamated) preference ordering of the majority of voters among the competing candidates may nevertheless be intransitive. A necessary condition for this to occur is that the various majorities are composed of different persons and there exist at least three candidates. Although we do not demonstrate this paradox in the Appendix, it may occur under all ranked voting procedures, as well as under the successive elimination procedure.

3.2.1.2 The Condorcet Winner Paradox (Condorcet 1785; Black 1958)

A candidate x is not elected despite the fact that it constitutes a "Condorcet Winner", i.e., despite the fact that x is preferred by a majority of the voters over each of the other competing alternatives.[2]

3.2.1.3 The Absolute Majority Paradox

This is a special case of the Condorcet winner paradox. A candidate x may not be elected despite the fact that it is the only candidate ranked first by an absolute majority of the voters.

[2] Fishburn (1974, p. 544) constructs an example with 101 voters and nine candidates two of whom are candidates a and w, such that w beats each of the other eight candidates by a (slim) majority of 51 to 50 (and hence is a Condorcet winner), whereas a beats each of the other seven candidates by a considerably larger majority. Fishburn states that "examples like this suggest that some cases which have a simple-majority [Condorcet] winner do *not* represent the most satisfactory social choice." We disagree with this statement and hold that a Condorcet winner, if one exists, ought *always* to be elected.

3.2.1.4 The Condorcet Loser or Borda Paradox (Borda 1784; Black 1958)

A candidate x is elected despite the fact that it constitutes a "Condorcet Loser" i.e., despite the fact that a majority of voters prefer each of the remaining candidates to x. This paradox is a special case of the violation of Smith's (1973) Condorcet principle. According to this principle, if it is possible to partition the set of candidates into two disjoint subsets, A and B, such that each candidate in A is preferred by a majority of the voters over each candidate in B, then no candidate in B ought to be elected unless all candidates in A are elected.

3.2.1.5 The Absolute Loser Paradox

This is a special case of the Condorcet loser paradox. A candidate x may be elected despite the fact that it is ranked last by a majority of voters.

3.2.1.6 The Pareto (or Dominated Candidate) Paradox (Fishburn 1974)

A candidate x may be elected while candidate y may not be elected despite the fact that *all* voters prefer candidate y to x.

3.2.2 Conditional Paradoxes

The nine best-known "conditional" paradoxes that may afflict voting procedures for electing a single candidate are the following:

3.2.2.1 Additional Support (or Lack of Monotonicity or Negative Responsiveness) Paradox (Smith 1973; Fishburn 1974a, Fishburn and Brams 1983)

If candidate x is elected under a given distribution of voters' preferences among the competing candidates, it is possible that, *ceteris paribus*, x may not be elected if some voter(s) *increase(s) his (their) support for* x by moving x to a higher position in his (their) preference ordering. Alternatively, if candidate x is not elected under a given distribution of voters' preferences among the competing candidates, it is possible that, *ceteris paribus*, x will be elected if some voter(s) *decrease(s)* his (their) support for x by moving x to a lower position in his (their) preference ordering.[3]

[3] Another version of the non-monotonicity paradox (which is not demonstrated in the Appendix) is a situation where x is elected in a given electorate but may not be elected if, *ceteris paribus*, additional voters join the electorate who rank x at the top of their preference ordering, or,

3.2.2.2 Reinforcement (or Inconsistency or Multiple Districts) Paradox (Young 1974)

If x is elected in each of several disjoint electorates, it is possible that, *ceteris paribus*, x will not be elected if all electorates are combined into a single electorate.

3.2.2.3 Truncation Paradox (Brams 1982; Fishburn and Brams 1983)

A voter may obtain a more preferable outcome if, *ceteris paribus*, he lists in his ballot only part of his (sincere) preference ordering among some of the competing candidates than listing his entire preference ordering among all the competing candidates.

3.2.2.4 No-Show paradox (Fishburn and Brams 1983; Ray 1986; Moulin 1988b; Holzman 1988/89; Pérez 1995)

This is an extreme version of the truncation paradox. A voter may obtain a more preferable outcome if he decides not to participate in an election than, *ceteris paribus*, if he decides to participate in the election and vote sincerely for his top preference(s).

3.2.2.5 Twin Paradox (Moulin 1988b)

This is a special version of the no-show paradox. Two voters having the same preference ordering may obtain a preferable outcome if, *ceteris paribus*, one of them decides not to participate in the election while the other votes sincerely.

3.2.2.6 Violation of the Subset Choice Condition (SCC) (Fishburn 1974b,c, 1977)

SCC requires that when there are at least three candidates and candidate x is the unique winner, then x must not become a loser whenever any of the original losers is removed and all other things remain the same. All the voting procedures discussed in this paper except the range voting (RV) and majority judgment (MJ) procedures violate SCC.[4] In the context of individual choice theory SCC is known as Chernoff's

alternatively, a situation where $x$is not elected in a given electorate but may be elected if, *ceteris paribus*, additional voters join the electorate who rank x at the bottom of their preference ordering.

[4]The RV and MJ procedures satisfy SCC because these procedures do not aggregate the individual voters' preference orderings into a social preference ordering in order to determine the winner. Under these procedures every candidate is ranked (on a cardinal or ordinal scale) by every voter,

condition (1954, p. 429, postulate 4) which states that if an alternative x chosen from a set T is an element of a subset S of T, then x must be chosen also from S.

3.2.2.7 Preference Inversion Paradox

If the individual preferences of each voter are inverted it is possible that, *ceteris paribus*, the (unique) original winner will still win.

3.2.2.8 Lack of Path Independence Paradox (Farquharson 1969; Plott 1973)

If the voting on the competing candidates is conducted sequentially rather than simultaneously, it is possible that candidate x will be elected under a particular sequence but not, *ceteris paribus*, under an alternative sequence.

3.2.2.9 Strategic Voting Paradox (Gibbard 1973; Satterthwaite 1975)

There are conditions under which a voter with full knowledge of how the other voters are to vote and the decision rule being used, would have an incentive to vote in a manner that does not reflect his true preferences among the competing alternatives. All known non-dictatorial voting procedures suffer from this paradox; it is not demonstrated in the Appendix.

3.3 Voting Procedures for Electing One out of Two or More Candidates

3.3.1 Non-ranked Voting Procedures

There are four main voting procedures for electing a single candidate where voters do not have to rank-order the candidates:

---

and the winner is that candidate whose average (or median) rank is highest. Thus the elimination of any losing candidate cannot affect, *ceteris paribus*, the identity of the original winner.

It may perhaps be assumed that under Approval Voting a voter will never vote for an alternative in a subset which s/he did not "approve" in the superset, and hence that Approval Voting, too, satisfies SCC. This assumption is debatable. It can easily be shown – as in Example 3.5.1.1. below – that when there are three alternatives among whom a voter has a linear preference ordering, it would always be rational for a voter under Approval Voting to vote for his/her second preference if his/her top preference is no longer available – even if originally s/he "approved" only of his/her top preference. By doing so s/he has nothing to lose but may obtain a better outcome than by abstaining – regardless of how all other voters are going to vote. Hence in our view Approval Voting may violate SCC.

3.3.1.1 Plurality (or First Past the Post, Briefly FPTP) Voting Procedure

This is the most common procedure for electing a single candidate, and is used, *inter alia*, for electing the members of the House of Commons in the UK and the members of the House of Representatives in the US. Under this procedure every voter casts one vote for a single candidate and the candidate obtaining the largest number of votes is elected.

3.3.1.2 Plurality with Runoff Voting Procedure

Under the usual version of this procedure up to two voting rounds are conducted. In the first round each voter casts one vote for a single candidate. In order to win in the first round a candidate must obtain either a special plurality (usually at least 40% of the votes) or an absolute majority of the votes. If no candidate is declared the winner in the first round then a second round is conducted. In this round only the two candidates who obtained the highest number of votes in the first round participate, and the one who obtains the majority of votes wins. This too is a very common procedure for electing a single candidate and is used, *inter alia*, for electing the President of France.

3.3.1.3 Approval Voting (Brams and Fishburn 1978, 1983)

Under this procedure every voter has a number of votes which is equal to the number of competing candidates, and every voter can cast one vote or no vote for every candidate. The candidate obtaining the largest number of votes is elected. So far this procedure has not been used in any public elections but is already used by several professional associations and universities in electing their officers.

3.3.1.4 Successive Elimination (Farquharson 1969)

This procedure is common in parliaments when voting on alternative versions of bills. According to this procedure voting is conducted in a series of rounds. In each round two alternatives compete; the one obtaining fewer votes is eliminated and the other competes in the next round against one of the alternatives which has not yet been eliminated. The alternative winning in the last round is the ultimate winner.

3.3.2 Ranked Voting Procedures That Are Not Condorcet-Consistent

Six ranked procedures under which every voter must rank-order all competing candidates – but which do not ensure the election of a Condorcet winner when one exists – have been proposed, as far as I know, during the last 250 years. These procedures are described below. Only one of these procedures (alternative vote) is used currently in public elections.

3.3.2.1 Borda's Count (Borda 1784; Black 1958)

This voting procedure was proposed by Jean Charles de Borda in a paper he delivered in 1770 before the French Royal Academy of Sciences entitled 'Memorandum on election by ballot' ('Mémoire sur les élections au scrutin'). According to Borda's procedure each candidate, x, is given a score equal to the number of pairs $(V,\ y)$ where V is a voter and y is a candidate such that V prefers x to y, and the candidate with the largest score is elected. Equivalently, each candidate x gets no points for each voter who ranks x last in his preference ordering, one point for each voter who ranks x second-to-last in his preference order, and so on, and $m-1$ points for each voter who ranks x first in his preference order (where m is the number of candidates). Thus if all n voters have linear preference orderings among the m candidates then the total number of points obtained by all candidates is equal to the number of voters multiplied by the number of paired comparisons, i.e., to $nm(m-1)/2$.

3.3.2.2 Alternative Vote (AV); (aka Instant Runoff Voting)

This is the version of the *single transferable vote* (STV) procedure (independently proposed by Carl George Andrae in Denmark in 1855 and by Thomas Hare in England in 1857) for electing a single candidate. It works as follows. In the first step one verifies whether there exists a candidate who is ranked first by an absolute majority of the voters. If such a candidate exists s/he is declared the winner. If no such candidate exists then, in the second step, the candidate who is ranked first by the smallest number of voters is deleted from all ballots and thereafter one again verifies whether there is now a candidate who is ranked first by an absolute majority of the voters. The elimination process continues in this way until a candidate who is ranked first by an absolute majority of the voters is found. The Alternative Vote procedure is used in electing the president of the Republic of Ireland, the Australian House of Representatives, as well as the mayors in some municipal elections in the US.

3.3.2.3 Coombs' Method (Coombs 1964, pp. 397–399; Straffin 1980; Coombs et al. 1984)

This procedure was proposed by the psychologist Clyde H. Coombs in 1964. It is similar to Alternative Vote except that the elimination in a given round under Coombs' method involves the candidate who is ranked last by the largest number of voters (instead of the candidate who is ranked first by the smallest number of voters under alternative vote).

3.3.2.4 Bucklin's Method (Hoag and Hallett 1926, pp. 485–491; Tideman 2006, p. 203)

This voting system can be used for single-member and multi-member districts. It is named after James W. Bucklin of Grand Junction, Colorado, who first promoted it in 1909. In 1913 the US Congress prescribed (in the Federal Reserve Act of 1913, Sect. 4) that this method be used for electing district directors of each Federal Reserve Bank.

Under Bucklin's method voters rank-order the competing candidates. The vote count starts like in the Alternative Vote method. If there exists a candidate who is ranked first by an absolute majority of the voters s/he is elected. Otherwise the number of voters who ranked every candidate in second place are added to the number of voters who ranked him/her first, and if now there exists a candidate supported by a majority of voters s/he is elected. If not, the counting process continues in this way by adding for each candidate his/her third, fourth, . . . rankings, until a candidate is found who is supported by an absolute majority of the voters. If two or more candidates are found to be supported by a majority of voters in the same counting round then the one supported by the largest majority is elected.[5]

3.3.2.5 Range Voting (Smith 2000)

According to this procedure the suitability (or level of performance) of every candidate is assessed by every voter and is assigned a (cardinal) grade (chosen from a pre-specified range) reflecting the candidate's suitability or level of performance in the eyes of the voter. The candidate with the highest average grade is the winner. This procedure is currently championed by Warren D. Smith (see http://rangevoting.org) and used to elect the winner in various sport competitions.

---

[5]However, it is unclear how a tie between two candidates, say a and b, ought to be broken under Bucklin's procedure when both a and b are supported in the same counting round by the same number of voters and this number constitutes a majority of the voters. If one tries to break the tie between a and b in such an eventuality by performing the next counting round in which all other candidates are also allowed to participate, then it is possible that the number of (cumulated) votes of another candidate, c, will exceed that of a and b.

To see this, consider the following simple example. Suppose there are 18 voters who must elect one candidate under Bucklin's procedure and whose preference orderings among four candidates, a, b, c, d are as follows: seven voters with preference ordering $a \succ b \succ c \succ d$, eight voters with preference ordering $b \succ a \succ c \succ d$, one voter with preference ordering $d \succ c \succ a \succ b$, and two voters with preference ordering $d \succ c \succ b \succ a$. None of the candidates constitutes the top preference of a majority of the voters. However, both a and b constitute the top or second preference by a majority of voters (15). If one tries to break the tie between a and b by performing the next (third) counting round in which c and d are also allowed to participate, then c will be elected (with 18 votes), but if only a and b are allowed to participate in this counting round then b will be elected (with 17 votes).

So which candidate ought to be elected in this example under Bucklin's procedure? As far as I know, Bucklin did not supply an answer to this question.

3.3.2.6 Majority Judgment (Balinski and Laraki 2007a,b, 2011)

According to this proposed procedure, the suitability (or level of performance) of every candidate is assessed by every voter and is assigned an ordinal grade (chosen from a pre-specified range) reflecting the candidate's suitability or level of performance in the eyes of the voter. The candidate with the highest median grade is the winner.

3.3.3 Ranked Voting Procedures that are Condorcet-consistent[6]

All the eight voting procedures described in this subsection require that voters rank-order all competing candidates. Under all these procedures a *Condorcet winner*, if one exists, is elected. The procedures differ from one another regarding which candidate gets elected when the social preference ordering contains a top cycle, i.e., when a Condorcet winner does not exist.

3.3.3.1 The Minimax Procedure

Condorcet specified that the Condorcet winner (whom he called 'the majority candidate') ought to be elected if one exists. However, according to Black (1958, pp. 174–175, 187) Condorcet did not specify clearly which candidate ought to be elected when the social preference ordering contains a top cycle. Black (1958, p. 175) suggests that "It would be most in accordance with the spirit of Condorcet's ... analysis ... to discard all candidates except those with the minimum number of majorities against them and then to deem the largest size of minority to be a majority, and so on, until one candidate had only actual or deemed majorities against each of the others." In other words, the procedure attributed by Black to Condorcet when cycles exist in the social preference ordering is a *minimax procedure*[7] since it chooses that candidate whose worst loss in the paired comparisons is the least bad. This procedure is also known in the literature as the *Simpson–Kramer rule* (see Simpson 1969; Kramer 1977).

---

[6] I list here only deterministic procedures. For a Condorcet-consistent probabilistic procedure see Felsenthal and Machover (1992). I also do not list here two Condorcet-consistent deterministic procedures proposed by Tideman (1987) and by Schultze (2003) because I do not consider satisfying (or violating) the independence-of-clones property, which is the main reason why these two procedures were proposed, to be associated with any voting paradox. (A phenomenon where candidate x is more likely to be elected when two clone candidates, y and y', exist, and where x is less likely to be elected when, *ceteris paribus*, one of the clone candidates withdraws, does not seem to me surprising or counter-intuitive).

[7] Young (1977, p. 349) prefers to call this procedure "The minimax function".

3.3.3.2 Dodgson's procedure (Black 1958, pp. 222–234; McLean and Urken, 1995, pp. 288–297)

This procedure is named after the Rev. Charles Lutwidge Dodgson, aka Lewis Carroll, who proposed it in 1876. It elects the Condorcet winner when one exists. If no Condorcet winner exists it elects that candidate who requires the fewest number of switches (i.e. inversions of two adjacent candidates) in the voters' preference orderings in order to make him the Condorcet winner.

3.3.3.3 Nanson's Method (Nanson 1883; McLean and Urken, 1995, ch. 14)

Nanson's method is a recursive elimination of Borda's method. In the first step one calculates for each candidate his Borda score. In the second step the candidates whose Borda score do not exceed the average Borda score of the candidates in the first step are eliminated from all ballots and a revised Borda score is computed for the uneliminated candidates. The elimination process is continued in this way until one candidate is left. If a (strong) Condorcet winner exists then Nanson's method elects him.[8]

3.3.3.4 Copeland's Method (Copeland 1951)

Every candidate x gets one point for every paired comparison with another candidate y in which an absolute majority of the voters prefer x to y, and half a point for every paired comparison in which the number of voters preferring x to y is equal to the number of voters preferring y to x. The candidate obtaining the largest sum of points is the winner.

3.3.3.5 Black's Method (Black 1958, p. 66)

According to this method one first performs all paired comparisons to verify whether a Condorcet winner exists. If such a winner exists then s/he is elected. Otherwise the winner according to Borda's count (see above) is elected.

[8]Although Nanson's procedure satisfies the strong Condorcet condition, i.e., it always elects a candidate who beats every other candidate in paired comparisons, this procedure may not satisfy the weak Condorcet condition which requires that if there exist(s) candidate(s) who is (are) unbeaten by any other candidate then this (these) candidate(s) – and only this (these) candidate(s) – ought to be elected. For an example of violation of the weak Condorcet condition by Nanson's procedure see Niou (1987).

3.3.3.6 Kemeny's Method (Kemeny 1959; Kemeny and Snell 1960; Young and Levenglick 1978; Young 1995)

Kemeny's method (aka *Kemeny–Young rule*) specifies that up to $m!$ possible social preference orderings should be examined (where m is the number of candidates) in order to determine which of these is the "most likely" true social preference ordering.[9] The selected "most likely" social preference ordering according to this method is the one where the number of pairs $(A, \ y)$, where A is a voter and y is a candidate such that A prefers x to y, and y is ranked below x in the social preference ordering is maximized. Given the voters' various preference orderings, Kemeny's procedure can also be viewed as finding the most likely (or the best predictor, or the best compromise) true social preference ordering, called the *median preference ordering*, i.e., that social preference ordering S that minimizes the sum, over all voters i, of the number of pairs of candidates that are ordered oppositely by S and by the ith voter.[10]

3.3.3.7 Schwartz's Method (Schwartz 1972; 1986)

Thomas Schwartz's method is based on the notion that a candidate x deserves to be listed ahead of another candidate y in the social preference ordering if and only if x beats or ties with some candidate that beats y, and x beats or ties with all candidates that y beats or ties with. The Schwartz set (from which the winner should be chosen) is the smallest set of candidates who are unbeatable by candidates outside the set. The Schwartz set is also called *GOCHA* (*Generalized Optimal Choice Axiom*).

3.3.3.8 Young's Method (Young 1977)

According to Fishburn's (1977, p. 473) informal description of Young's procedure "[it] is like Dodgson's in the sense that it is based on altered profiles that have candidates who lose to no other candidate under simple majority. But unlike

---

[9] Tideman (2006, pp. 187–189) proposes two heuristic procedures that simplify the need to examine all $m!$ preference orderings.

[10] According to Kemeny (1959) the distance between two preference orderings, R and R', is the number of pairs of candidates (alternatives) on which they differ. For example, if $R = a \succ b \succ c \succ d$ and $R' = d \succ a \succ b \succ c$, then the distance between R and R' is 3, because they agree on three pairs $[(a \succ b), (a \succ c), (b \succ c)]$ but differ on the remaining three pairs, i.e., on the preference ordering between a and d, b and d, and between c and d. Similarly, if R'' is $c \succ d \succ a \succ b$ then the distance between R and R'' is 4 and the distance between R' and R'' is 3. According to Kemeny's procedure the most likely social preference ordering is that R such that the sum of distances of the voters' preference orderings from R is minimized. Because this R has the properties of the median central measure in statistics it is called the *median preference ordering*. The median preference ordering (but not the *mean preference ordering* which is that R which minimizes the sum of the squared differences between R and the voters' preference orderings) will be identical to the possible social preference ordering W which maximizes the sum of voters that agree with all paired comparisons implied by W.

Dodgson, Young deletes voters rather than inverting preferences to obtain the altered profiles. His procedure suggests that we remain most faithful to Condorcet's principle if the choice set consists of alternatives that can become simple majority nonlosers with removal of the fewest number of voters."

3.4 Summary

As can be seen from Tables 3.1–3.3, seven procedures (Alternative Vote, Coombs, Bucklin, Majority Judgment, Minimax, Dodgson, and Young) are susceptible to the largest number of paradoxes (10), whereas the plurality (first-past-the-post) and Borda's procedures are susceptible to the smallest number of paradoxes (6).

Of the nine Condorcet-consistent procedures, six procedures (successive elimination, minimax, Dodgson's, Nanson's, Schwartz's, and Young's) are dominated by the other three procedures (Black's, Copeland's and Kemeny's) in terms of the paradoxes to which these procedures are susceptible.

However, the number of paradoxes to which each of the various voting procedures surveyed here is vulnerable may be regarded as meaningless or even misleading. This is so for two reasons.

Table 3.1 Susceptibility of non-ranked procedures to voting paradoxes

Procedure	Plurality	Plurality ϖ runoff	Approval voting	Successive elimination
Paradox				
Condorcet pdx (cyclical majorities)	−	−	−	+
Condorcet winner pdx	+	+	+	−
Absolute majority pdx	−	−	⊕	−
Condorcet loser pdx	⊕	−	+	−
Absolute loser pdx	⊕	−	⊕	−
Pareto dominated candidate	−	−	⊕	⊕
Lack of monotonicity	−	⊕	−	−
Reinforcement	−	+	−	+
No-show	−	+	−	+
Twin	−	+	−	+
Truncation	−	−	−	+
Subset choice condition (SCC)	+	+	+	+
Preference inversion	+	+	+	−
Path independence	−	−	−	+
Strategic voting	+	+	+	+
Total ⊕ signs	2	1	4	1
Total + & ⊕ signs	6	8	8	9

Notes:
A + sign indicates that a procedure is vulnerable to the specified paradox
A ⊕ sign indicates that a procedure is vulnerable to the specified paradox which seems to us an especially intolerable paradox
A − sign indicates that a procedure is not vulnerable to the specified paradox
It is assumed that all voters have linear preference ordering among all competing candidates

Table 3.2 Susceptibility of ranked non Condorcet-consistent procedures to voting paradoxes

Procedure	Borda	Alternative Vote (AV) STV	Coombs	Bucklin	Range Voting	Majority Judgment
Paradox						
Condorcet pdx (cyclical majorities)	+	+	+	+	+	+
Condorcet winner pdx	+	+	+	+	+	+
Absolute majority pdx	⊕	−	−	−	⊕	⊕
Condorcet loser pdx	−	−	−	⊕	⊕	⊕
Absolute loser pdx	−	−	−	−	⊕	⊕
Pareto dominated candidate	−	−	−	−	−	−
Lack of monotonicity	−	⊕	⊕	−	−	−
Reinforcement	−	+	+	+	−	+
No-show	−	+	+	+	−	+
Twin	−	+	+	+	−	+
Truncation	+	+	+	+	+	+
Subset choice condition (SCC)	+	+	+	+	−	−
Preference inversion	−	+	+	+	−	−
Path independence	−	−	−	−	−	−
Strategic voting	+	+	+	+	+	+
Total ⊕ signs	1	1	1	1	3	3
Total + and ⊕ signs	6	10	10	10	7	10

Notes:
A + sign indicates that a procedure is vulnerable to the specified paradox
A ⊕ sign indicates that a procedure is vulnerable to the specified paradox which seems to us an especially intolerable paradox
A − sign indicates that a procedure is not vulnerable to the specified paradox
It is assumed that all voters have linear preference ordering among all competing candidates

First, some paradoxes are but special cases of other paradoxes or may induce the occurrence of other paradoxes, as follows:

- A procedure which is vulnerable to the absolute majority paradox is also vulnerable to the Condorcet winner paradox;
- A procedure which is vulnerable to the absolute loser paradox is also vulnerable to the Condorcet loser paradox;
- Except for the range voting and majority judgment procedures, all procedures surveyed in this chapter that are vulnerable to the Condorcet loser paradox are also vulnerable to the preference inversion paradox.
- The five procedures surveyed in this chapter which may display lack of monotonicity are also susceptible to the No-Show paradox[11];

[11] Campbell and Kelly (2002) devised a non-monotonic voting rule that does not exhibit the No-Show paradox. However, as this method violates the anonymity and neutrality conditions and hence has not been considered seriously for actual use, we ignore it.

Table 3.3 Susceptibility of ranked Condorcet-consistent procedures to voting paradoxes

Procedure	Minimax	Dodgson	Black	Copeland	Kemeny	Nanson	Schwartz	Young
Paradox								
Condorcet pdx (cyclical majorities)	+	+	+	+	+	+	+	+
Condorcet winner pdx	−	−	−	−	−	−	−	−
Absolute majority pdx	−	−	−	−	−	−	−	−
Condorcet loser pdx	⊕	⊕	−	−	−	−	−	⊕
Absolute loser pdx	⊕	−	−	−	−	−	−	⊕
Pareto dominated cand.	−	−	−	−	−	−	⊕	−
Lack of monotonicity	−	⊕	−	−	−	⊕	−	−
Reinforcement	+	+	+	+	+	+	+	+
No-show	+	+	+	+	+	+	+	+
Twin	+	+	+	+	+	+	+	+
Truncation	+	+	+	+	+	+	+	+
SCC	+	+	+	+	+	+	+	+
Preference inversion	+	+	−	−	−	−	−	+
Path independence	−	−	−	−	−	−	−	−
Strategic voting	+	+	+	+	+	+	+	+
Total ⊕ signs	2	2	0	0	0	1	1	2
Total + & ⊕ signs	10	10	7	7	7	8	8	10

Notes:
A + sign indicates that a procedure is vulnerable to the specified paradox
A ⊕ sign indicates that a procedure is vulnerable to the specified paradox which seems to us an especially intolerable paradox
A − sign indicates that a procedure is not vulnerable to the specified paradox
It is assumed that all voters have linear preference ordering among all competing candidates

- All Condorcet-consistent procedures are susceptible to the no-show paradox and hence also to the twin paradox when there exist at least four candidates.[12]

Second, and more importantly, not all the surveyed paradoxes are equally undesirable. Although assessing the severity of the various paradoxes is largely a subjective matter, *there seems to be a wide consensus that a voting procedure which is susceptible to an especially serious paradox (denoted by ⊕ in Tables 3.1–3.3), i.e., a voting procedure which may elect a pareto-dominated candidate, or elect a Condorcet (and absolute) loser, or display lack of monotonicity, or not elect an absolute winner, should be disqualified as a reasonable voting procedure regardless of the probability that these paradoxes may occur*. On the other hand, the degree of severity that should be assigned to the remaining paradoxes should depend, *inter alia*, on the likelihood of their occurrence under the procedures that are vulnerable

[12]Although all Condorcet-consistent procedures are also susceptible to the Reinforcement paradox, there is no logical connection between this paradox and the no-show paradox. As mentioned by Moulin (1988b, pp. 54–55), when there are no more than three candidates there exist Condorcet-consistent procedures which are immune to both the no-show and twin paradoxes, e.g., the minimax procedure which elects the candidate to whom the smallest majority objects.

to them. Thus, for example, a 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 procedures – should be deemed more desirable (and the paradoxes it may display more tolerable) than a procedure which can display the same paradox when a Condorcet winner exists.[13]

Additional criteria which should be used in assessing the relative desirability of a voting procedure are what may be called *administrative-technical criteria*. The main criteria belonging to this category are the following:

– *Requirements from the voter*: some voting procedures make it more difficult for the voter to participate in an election by requiring him/her to rank-order all competing candidates, whereas other procedures make it easier for the voter by requiring him/her to vote for just one candidate or for any candidate(s) s/he approves.
– *Ease of understanding how the winner is selected*: In order to encourage voters to participate in an election a voting procedure must be transparent, i.e., voters must understand how their votes (preferences) are aggregated into a social choice. Thus a voting procedure where the winner is the candidate who received the plurality of votes is easier to explain – and considered more transparent – than a procedure which may involve considerable mathematical calculations (e.g., Kemeny's) in order to determine the winner.
– *Ease of executing the elections*: Election procedures requiring only one voting (or counting) round are more easily executed than election procedures that may require more than one voting (or counting) round. Similarly, election procedures requiring to count only the number of votes received by each candidate are easier to conduct than those requiring the conduct of all $m(m-1)/2$ paired contests between all m candidates, or those requiring the examination of up to $m!$ possible social preference orderings in order to determine the winner.
– *Minimization of the temptation to vote insincerely*: Although all voting procedures are vulnerable to manipulation, i.e., to the phenomenon where some voters may benefit if they vote insincerely, some voting procedures (e.g., Borda's count, Range voting) are susceptible to this considerably more than others.
– *Discriminability*: One should prefer a voting procedure which is more discriminate, i.e., it is more likely to select (deterministically) a unique winner than produce a set of tied candidates – in which case the employment of additional means are needed to obtain a unique winner. Thus, for example, when the social preference ordering is cyclical then, *ceteris paribus*, Schwartz's and Copeland's methods are considerably less discriminating than the remaining Condorcet-consistent procedures surveyed in this chapter.

[13]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 compared procedures. Such knowledge is still lacking with respect to most voting procedures and paradoxes.

Of course there may exist conflicts between some of these technical-administrative criteria. For example, a procedure like Kemeny's which, on the one hand, is more difficult to execute in practice and to explain to prospective voters (and hence less transparent), is, on the other hand, more discriminate and less vulnerable to insincere behavior.

So in view of all the above criteria, which of the 18 surveyed voting procedures do I think should be preferred? Since the weakest extension of the majority rule principle when there are more than two candidates is the Condorcet winner principle, I think that the electoral system which ought to be used for electing one out of $m \geq 2$ candidates should be Condorcet-consistent.

But as one does not know before an election is conducted whether a Condorcet winner will exist or whether the social preference ordering will contain a top cycle, which of the nine Condorcet-consistent procedures surveyed and exemplified in this paper should be preferred in case a top cycle exists? In this case I think that the Successive Elimination procedure and Schwartz's procedure should be readily disqualified because of their vulnerability to electing a pareto-dominated candidate, Dodgson's and Nanson's procedures should be readily disqualified because of their lack of monotonicity, and the minimax and Young's procedures should be readily disqualified because of their vulnerability to electing an absolute or a Condorcet loser. Although Black's procedure cannot elect a Condorcet loser, it may nevertheless come quite close to it because, as demonstrated in Example 3.5.13.3 below, it violates Smith's (1973) Condorcet principle, so this procedure too seems to me not considerably more desirable than the minimax and Young's procedures.

This leaves us with a choice between the remaining two Condorcet-consistent procedures – Copeland's and Kemeny's. The choice between them depends on the importance one assigns to the above-mentioned technical-administrative criteria. Both these procedures require voters to rank-order all candidates. However, Copeland's method is probably easier than Kemeny's to explain to lay voters, as well as, when the number of candidates is large, may involve considerably fewer calculations in determining who is (are) the ultimate winner(s). Kemeny's procedure, on the other hand, is more discriminate than Copeland's when the number of candidates is relatively small, and is probably also – because of its increased complexity in determining the ultimate winner – less vulnerable to insincere voting. So if I would have to choose between these two procedures I would choose Kemeny's because most elections where a single candidate must be elected usually involve relatively few contestants – in which case Kemeny's procedure seems to have an advantage over Copeland's procedure. Moreover, as I mentioned in the description of Kemeny's procedure and as argued by Young (1995, pp. 60–62), Kemeny's procedure has also the advantage that it can be justified not only from Condorcet's perspective of the maximum likelihood rule, but also as choosing for the entire society the "median preference ordering" – which can be viewed from the perspective of modern statistics as the best compromise between the various rankings reported by the voters.

Chapter 13
And the Loser Is... Plurality Voting

Jean-François Laslier

13.1 Introduction

Experts have different opinions as to which is the best voting procedure. The Leverhulme Trust sponsored 2010 *Voting Power in Practice* workshop, held at the Chateau du Baffy, Normandy, from 30 July to 2 August 2010, was organized for the purpose of discussing this matter. Participants of the workshop were specialists in voting procedures and, during the wrap-up session at the end of the workshop, it was decided to organize a vote among the participants to elect "the best voting procedure". The present paper reports on this vote. It contains in the Appendix statements by some of the voters/participants about this vote and voting rules in general.

13.2 The Vote

Previous discussion had shown that different voting rules might be advisable under different circumstances, so that a more concrete problem than "What is the best voting rule" should be tackled. The question for the vote was: "What is the best voting rule for your town to use to elect the mayor?"

Even with this phrasing, it was realized afterwards that not all participants had exactly the same thing in mind. In particular, some of them were thinking of a large electorate and some were rather thinking of a committee (the city council) as the electorate. This can be inferred from the participants' comments in the Appendix and is clearly a weakness of this "experiment."

J.-F. Laslier (✉)
CNRS and Department of Economics, Ecole Polytechnique, Palaiseau, France
e-mail: jean-francois.laslier@polytechnique.edu

D.S. Felsenthal and M. Machover, *Electoral Systems*, Studies in Choice and Welfare,
DOI 10.1007/978-3-642-20441-8_13,

Of course, an interesting feature of this vote is the fact that it was a vote on voting rules *by voting theorists*. So the participants arrived with quite a heavy background of personal knowledge and ideas. But the way the vote was improvised was such that no one had much time to think things over, discuss and coordinate with others, or calculate. Moreover, no candidates were clear common knowledge front-runners, and the final result was apparently not anticipated by most voters.

The possibilities of strategic manipulation were thus quite limited and one can indeed see from the comments that most of these approval votes should be interpreted as the expression of sincere individual opinions. As one referee pointed out: this vote may be the last "naive" vote on voting rules. This adds a particular significance to its result, and also suggests that the experiment should be done again, now that the results are known.[1]

13.2.1 Candidates: The Voting Rules in Question

The set of "candidates," that is the list of considered voting rules was rather informally decided: participants just wrote on the paper board voting rules to be voted upon. Eighteen voting rules were nominated, the definitions of which can be found in Appendix B.

Some rules should really be considered as possible ways to organize elections and will, in usual circumstances, provide indeed a unique winner. Others will often yield not a single winner but a set of possible winners, among which the final choice has to be made by one means or another. The list contains several Condorcet-consistent rules which agree on a unique outcome when the Condorcet winner exists[2] but which differ when there is no Condorcet winner and, in that case, often yield several winners. For instance the Uncovered set is a singleton only if there is a Condorcet winner,[3] the Copeland winner is always in the Uncovered set and the Uncovered set is always included in the Top Cycle. The voting rules also differ as to their informational basis.

1. Some of them require very little information: Plurality voting and Majority voting with a runoff simply ask the voter to provide the name of one (or two) candidates. Approval Voting asks the voter to say "yes" or "no" to each candidate.
2. Most rules require that the voter ranks the candidates: this is the classical framework of Arrowian social choice (Arrow 1951). There is no inter-personal comparisons of alternatives, which means that the ballots are not intended to convey interpretation of the kind "candidate a is better for voter i than for

[1] My guess, based on the theoretical analysis of strategic voting under Approval Voting, is that the result would not be different.

[2] The existence of several Condorcet winners simultaneously is a rare phenomenon.

[3] No randomization scheme was considered. In particular the optimal solutions to the Condorcet paradox studied by Laffond et al. (1993) and Dutta and Laslier (1999) were not on the list of voting procedures.

voter j". The intra-personal structure is purely ordinal, which means that we may know that a is better than b for i, but we cannot know *how much* better.

3. Finally, some of them allow inter-personal comparisons, with intra-personal comparisons being ordinal (Leximin, Majority Judgement) or cardinal (Range Voting).

Most rules extend the majority principle in the sense that, if there are only two candidates, they select the one preferred by a majority of the voters. Range voting, which maximizes the average evaluation, does not fulfill this principle: indeed, according to classical utilitarianism, if a majority of voters slightly prefer a to b while a minority strongly prefers b to a, it may be better to choose b than a, against the majority principle. Therefore, under Range voting, if voters reflect in their vote this pattern of interpersonal comparisons, the minority candidate b may well be elected against a. Such is also the case for the "Majority judgement" system (despite its name) which maximizes the median evaluation and for the Leximin, which maximizes the worst evaluation.

13.2.2 The Procedure

13.2.2.1 The Electorate

The 22 voters were the participants of the workshop (one participant abstained). Some of them are advocates of a specific voting rule; for instance, Ken Ritchie and Alessandro Gardini are active in Great Britain in promoting the "Alternative Vote": a system of vote transfers also known as the "Hare" system. Others, like Dan Felsenthal, are advocates of the Condorcet principle and strongly defended this principle during the workshop. But it is fair to say that most of the participants would say that different voting rules have advantages and disadvantages. This might be one of the reasons why no-one objected to the use of Approval voting for this particular vote.

13.2.2.2 The Voting Rule

We used Approval voting for this election. Somebody made the suggestion and there was no counter-proposal. In retrospect, this choice was quite natural: this procedure is fast and easy to use even if the number of candidates is large. Asking voters to rank the 18 candidates was hardly feasible in our case. Approval voting is also advisable when the set of alternatives has been loosely designed and contains very similar candidates.[4] One may nevertheless regret that we lost the occasion to gather, through the vote, more information on the participants' opinions about the different

[4]See the cloning-consistency condition (Tideman 1987) and the composition-consistency property (Laffond et al. 1996).

Table 13.1 Number of approved candidates

Number of approvals	0	1	2	3	4	5	6	7	8	9	10	> 10	Total
Number of ballots	0	2	7	3	5	2	1	1	0	0	1	0	22

Table 13.2 Approval scores

Voting rule		Approvals	Approving percentage
Approval voting	`App`	15	68.18
Alternative vote	`Alt`	10	45.45
Copeland	`Cop`	9	40.91
Kemeny	`Kem`	8	36.36
Two-round majority	`2R`	6	27.27
Coombs	`Coo`	6	27.27
Simpson	`Sim`	5	22.73
Majority judgement	`Bal`	5	22.73
Borda	`Bor`	4	18.18
Black	`Bla`	3	13.64
Range voting	`RV`	2	9.09
Nanson	`Nan`	2	9.09
Leximin	`Lex`	1	4.54
Top-cycle	`TC`	1	4.54
Uncovered set	`UC`	1	4.54
Fishburn		0	0
Untrapped set		0	0
Plurality		0	0

voting rules. Hopefully the next section, where results are presented, will show that we can already learn quite a lot from the analysis of the Approval ballots.

13.3 The Results

13.3.1 Approval Score and Other Indicators

Voters approved on average 3.55 candidates out of 18, with a distribution provided in Table 13.1. This figure is not at odds with what has been observed in other circumstances (Laslier and Sanver 2010).

Table 13.2 provides the scores of the candidates:

Approvals. This is the number of voters who approve the candidate.
Approval score. This is the percentage of the population who approve the candidate. *Approval Voting* is approved by 15 voters out of 22, that is 68.18%.

Table 13.3 Various indicators

	Approvals	Markov	Focus	Central.	Simil.	Satisf.	Dilution
`App`	15	36.44	12.53	15.65	10.44	5.14	4.07
`Alt`	10	11.70	9.01	11.90	8.39	2.67	4.50
`Cop`	9	14.60	6.90	9.86	8.12	3.24	4.22
`Kem`	8	10.56	6.56	8.16	7.61	2.39	4.00
`2R`	6	6.91	5.50	7.14	6.86	1.71	4.5
`Coo`	6	6.28	4.88	7.48	6.89	1.63	4.67
`Sim`	5	2.82	7.17	8.50	6.52	0.99	6.00
`Bal`	5	2.26	4.69	7.48	6.62	1.05	5.40
`Bor`	4	4.42	5.08	5.78	6.16	1.24	5.25
`Bla`	3	0.86	5.59	6.12	5.83	0.47	7.00
`RV`	2	1.95	1.60	1.70	5.55	0.70	3.50
`Nan`	2	0.29	4.35	5.10	5.49	0.24	8.50
`Lex`	1	0.42	1.67	1.36	5.17	0.20	5.00
`TC`	1	0.30	1.82	1.70	5.17	0.17	6.00
`UC`	1	0.21	2.25	2.04	5.16	0.14	7.00

Approval Voting is the winner of the election. It is worth noticing that it is the only candidate approved by more than half of the voters.[5] Three candidates received no vote at all: *Fishburn*, *Untrapped Set*, and *Plurality*.

There are actually many different ways to compute scores and other indicators from a set of Approval ballots. Table 13.3 provides some, which are now defined. The number of voters who approved of both candidates c and c' is called the association of c and c', and is denoted by $as(c, c')$. The number of voters who approved c is denoted by $as(c)$.

Markov score. This score is computed as follows. The candidate "present at date t" is denoted $c(t)$. At date t chose at random one voter v. If v approves $c(t)$, keep this candidate for the next date: $c(t+1) = c(t)$. If not choose $c(t+1)$ at random among the candidates that v approves. This defines a Markov chain over candidates whose stationary distribution is the Markov score. For instance a candidate with Markov score 0.3 is, in the long run of this process, present 30% of the time.

Focus. The focus of candidate c is the sum over all candidates k of the fraction of k-voters who also approved c.

$$f(c) = \sum_k \frac{as(c,k)}{as(k)}.$$

The focus measures the ability of a candidate to attract votes from voters who also voted for others.

[5]One voter wrote on his/her ballot "Approval Voting with a runoff." This procedure was not on the list. This ballot was counted as an approbation of Approval voting.

Centrality. This indicator is based on the following Markov chain. The transition probability from c to c' is $as(c, c')/ \sum_{c \neq c'} as(c, c')$. The centrality measure is the associated stationary probability. This is a natural measure of centrality in the multi-graph where there is a link between two candidates each time a voter approves them both.

Similarity. This indicator is based on the following Markov chain. Given the candidate c, one chooses at random a voter v. If v approves c one replaces c by c' chosen at random among the candidates that v approves. If v does not approve c, one replaces c by c' chosen at random among all the candidates. The similarity measure is the associated stationary probability. This means that, given a candidate c, one looks for a candidate c' which is similar to c in the sense that a voter has approved both.

Satisfaction. If v has approved $B(v)$ candidates, count $1/B(v)$ points for each. The total count of candidate c is thus between 0 and the number of voters, and the sum over candidates is the number of voters. (See Kilgour 2010.)

Dilution. This is the average number of candidates approved by the voters who approve a given candidate. Let $as(v, c, k)$ be 1 if voter v approves both c and k, and 0 ifnot. Then:

$$dil(c) = \frac{1}{as(c)} \sum_v \sum_k as(v, c, k).$$

Notice that this indicator can be computed with a formula somehow dual to the focus:

$$dil(c) = \sum_k \frac{as(c, k)}{as(c)}.$$

The dilution thus measures to what extent supporters of a candidate also vote for other candidates. It should not be interpreted as an indicator of the strength of the candidate but as a part of the description of the electorate of the candidate: do these voters give exclusive support (low dilution), or do they support many other candidates (high dilution).

These indicators are all highly correlated with the approval score, except for the dilution (see Table 13.4).

Table 13.4 Correlations among indicators

	Approvals	Markov	Focus	Central.	Simil.	Adjust.	Dilution
Approvals	1	0.935	0.930	0.958	0.999	0.980	−0.547
Markov	0.935	1	0.851	0.981	0.930	0.878	−0.285
Focus	0.930	0.851	1	0.839	0.943	0.966	−0.511
Central	0.958	0.981	0.839	1	0.952	0.897	−0.347
Simil.	0.999	0.930	0.943	0.952	1	0.984	−0.546
Adjust.	0.980	0.878	0.966	0.897	0.984	1	−0.603
Dilution	−0.547	−0.285	−0.511	−0.347	−0.546	−0.603	1

13.3.2 Structure of the Set of Candidates

Table 13.5 shows the number of voters who approved each pair of candidates, and Table 13.6 shows the distribution of these association numbers for each candidate. For instance $7/10 = 70\%$ of the voters who approved the *Alternative Vote* also approved *Approval Voting* while $7/15 = 47\%$ of the voters who approved *Approval Voting* also approved the *Alternative Vote*. It is interesting to note that 83% of the supporters of two-round majority voting also support the *Alternative Vote*, but such is the case of only 22% of the *Copeland* supporters. One may also notice that all the supporters of the *Majority Judgement* are also supporters of *Approval Voting*.

To obtain a more global view, one may compute various distances between candidates. Consider for instance for the similarity index

$$sim(c,c') = \frac{as(c,c')}{as(c)} + \frac{as(c,c')}{as(c')}$$

which ranges from 0 (when the electorates of c and c' are disjoint) to 2 (when they are identical) and define

$$dist(c,c') = 2 - sim(c,c').$$

It turns out that there exists a very good Euclidean representation of the 15 candidates in 3 dimensions, that renders 90% of the sum of square of distances.[6] Figures 13.1 and 13.2 are side views of this representation. *Approval Voting* is in the center. The points on the right are rules which are important in the social choice literature: *Uncovered set*, *Copeland*, *Nanson*, *Kemeny*, *Simpson*, even if they are not very practical. *Borda* is in this group, close to *Nanson*. The points on the left contain three practical solutions to the voting problem: *Two-round majority*, the *Alternative Vote*, and *Black*. *Leximin* is not far from this group. *Coombs* and *Majority Judgement* are close one to the other, with *Range Voting* not far. The *Top-cycle* is isolated.

This structure reflects the vote profile since by definition, two voting rules are represented close one to the other when the same voters approved both.

Studying how candidate rules are associated in the voters' ballots, it appears that the winner is receiving votes associated with all the other candidates. *Approval Voting* can be described as a "centrist" candidate in this vote. Even if one can detect some pattern in the vote profile that differentiates votes for more "theoretical" rules from votes for more "practical" rules, the electorate does not appear to be split.

[6] See Appendix C for more details.

Table 13.5 Association matrix

	App	Alt	Cop	Kem	2R	Coo	Sim	Bal	Bor	Bla	RV	Nan	Lex	TC	UC
App	**15**	7	7	4	3	3	4	5	3	3	2	2	1	1	1
Alt	7	**10**	2	3	5	4	3	3	1	3	1	1	1	1	0
Cop	7	2	**9**	4	1	2	3	4	2	1	0	2	0	0	1
Kem	4	3	4	**8**	1	3	3	1	1	1	0	1	0	1	1
2R	3	5	1	1	**6**	3	1	1	2	2	0	1	1	0	0
Coo	3	4	2	3	3	**6**	1	2	1	1	1	1	0	0	0
Sim	4	3	3	3	1	1	**5**	2	2	2	0	2	0	1	1
Bal	5	3	4	1	1	2	2	**5**	1	1	1	1	0	0	0
Bor	3	1	2	1	2	1	2	1	**4**	1	0	2	0	0	1
Bla	3	3	1	1	2	1	2	1	1	**3**	0	1	1	1	0
RV	2	1	0	0	0	1	0	1	0	0	**2**	0	0	0	0
Nan	2	1	2	1	1	1	2	1	2	1	0	**2**	0	0	1
Lex	1	1	0	0	1	0	0	0	0	1	0	0	**1**	0	0
TC	1	1	0	1	0	0	1	0	0	1	0	0	0	**1**	0
UC	1	0	1	1	0	0	1	0	1	0	0	1	0	0	**1**

Table 13.6 Conditional association matrix

	App	Alt	Cop	Kem	2R	Coo	Sim	Bal	Bor	Bla	RV	Nan	Lex	TC	UC
App	**1**	0.7	0.78	0.5	0.5	0.5	0.8	1	0.75	1	1	1	1	1	1
Alt	0.47	**1**	0.22	0.38	0.83	0.67	0.6	0.6	0.25	1	0.5	0.5	1	1	0
Cop	0.47	0.2	**1**	0.5	0.17	0.33	0.6	0.8	0.50	0.33	0	1	0	0	1
Kem	0.27	0.3	0.44	**1**	0.17	0.5	0.6	0.2	0.25	0.33	0	0.5	0	1	1
2R	0.2	0.5	0.11	0.12	**1**	0.5	0.2	0.2	0.5	0.67	0	0.5	1	0	0
Coo	0.2	0.4	0.22	0.38	0.5	**1**	0.2	0.4	0.25	0.33	0.5	0.5	0	0	0
Sim	0.27	0.3	0.33	0.38	0.17	0.17	**1**	0.4	0.50	0.67	0	1	0	1	1
Bal	0.33	0.3	0.44	0.12	0.17	0.33	0.4	**1**	0.25	0.33	0.5	0.5	0	0	0
Bor	0.20	0.1	0.22	0.12	0.33	0.17	0.4	0.2	**1**	0.33	0	1	0	0	1
Bla	0.20	0.3	0.11	0.12	0.33	0.17	0.4	0.2	0.25	**1**	0	0.5	1	1	0
RV	0.13	0.1	0	0	0	0.17	0	0.2	0	0	**1**	0	0	0	0
Nan	0.13	0.1	0.22	0.12	0.17	0.17	0.4	0.2	0.5	0.33	0	**1**	0	0	1
Lex	0.07	0.1	0	0	0.17	0	0	0	0	0.33	0	0	**1**	0	0
TC	0.07	0.1	0	0.12	0	0	0.2	0	0	0.33	0	0	0	**1**	0
UC	0.07	0	0.11	0.12	0	0	0.2	0	0.25	0	0	0.5	0	0	**1**

Fig. 13.1 Axes 1 and 2

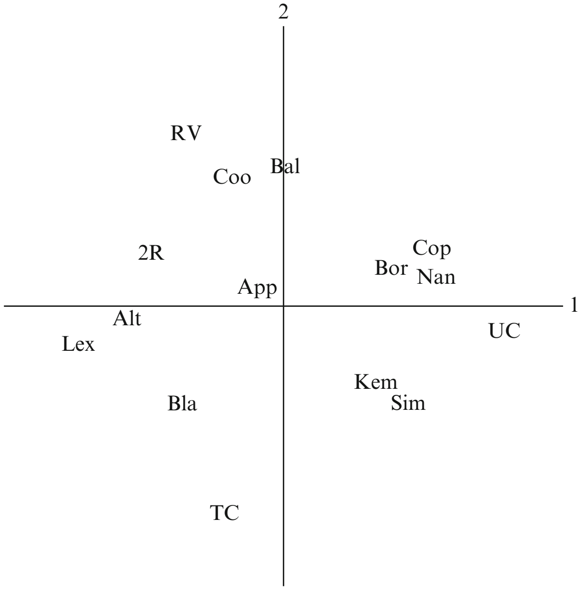
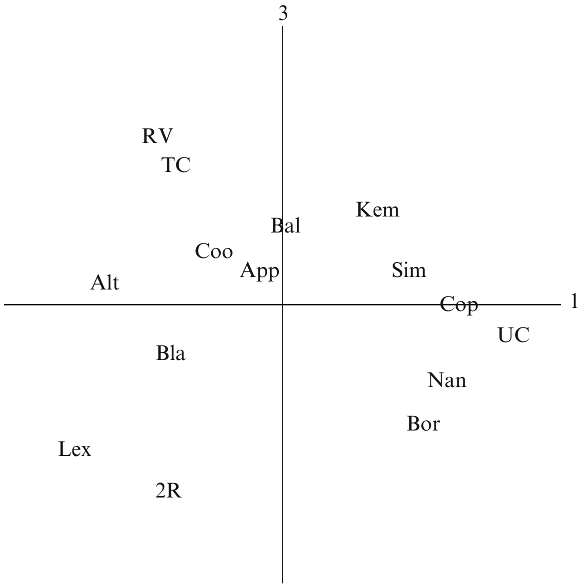


Fig. 13.2 Axes 1 and 3



13.4 Conclusion

The analysis of the approval ballots show that *Approval Voting* was a clear winner of this election. This is somehow surprising since *Approval Voting* was not much discussed during the workshop. But this voting rule has already received a lot of attention in the academic literature, and was certainly familiar to the participants.

A striking fact is that *Plurality* rule (First Past the Post) received no approval. The whole ranking of the candidate procedures according to their approval scores seems also robust since alternative ways to count the ballots produce rather similar rankings.

Appendix A: Contributions of the Participants

Participants were ex post invited to write a brief statement that would explain their vote and their view about this "election". One-half of them did it, after having read my own contribution (see below) as an example.

It must be acknowledged that, from a scientific point of view, the experimental protocol was a little loose. The nomination procedure for the set of candidate voting rules was informal and voters did not have all the time required to learn everything about all of them.

> Fuad Aleskerov: "From my point of view our voting for the rules was rather spontaneous. For instance, I know more than 30 rules which can be listed for our voting after reading careful studies, by colleagues, of their properties."

It is also worth remarking that, under Approval Voting, it is not so easy to remember after several days or weeks which candidates you approved out of 18.

> Marc Kilgour: "I really can't remember very well why I voted the way I did. As I recall, the objective was to propose a system to elect the mayor of a town, without any indication of the number of candidates. I think I assumed that there would not be many. I followed the approval strategy of approving everything that seemed to be above average "utility," whatever that would mean. I voted for approval (my actual favorite) and range because they focus on acceptability rather than ranks. I also voted for two or three others that, it seemed to me, were complicated enough to be likely to produce something that would maximize the sum of the utilities but at the same time sophisticated enough to avoid features I don't like such as non-monotonicity. Beyond that, I can't remember much."

This point holds certainly true, as well, for preference-based balloting: it is not so easy to remember how you ranked the whole set of alternatives; it distinguishes these systems from the familiar single-name Plurality and Two-Round Majority voting rules. The set of received contributions, which follow, show well the variety of view points of the experts in the field.

A.1 A. Baujard

The question we have been asked implied to apply the rule in a real political context, involving citizens with their own desires, their own intellectual abilities and the differences among them. Choosing the good rule for a democratic mayor election predicates to pay attention to these features rather than to my own preferences among rules. Through the field-experiments we have conducted, I have learnt that most voters are frustrated by voting rules which gives little scope to the expression of their nuanced preferences, which are sometimes tinted by hesitations, indifference, significant differences between strong vs. weak preferences. These nuances have yet a strong impact on results especially in the case of uninominal voting rules, where a clear cut decision is always required to infer individual preferences. In a democracy, I claim that this should question the legitimacy of the winner in such context. I am therefore not convinced by plurality rules, whatever one or two rounds, Condorcet principle, or any uninominal voting rules in general.

Among plurinominal rules, I focus on the importance of simplicity to explain the rule, and to vote. Above all, I paid attention to transparency, meaning a wide understanding of the process of deriving a result from the ballots, and the ability of citizens to take actively part in the process of counting the votes. These desired properties rule out Alternative Vote, Borda rules and Majority Judgement among others.

Many voters in our experiments spontaneously preferred range voting, begging for the ability of giving negative grades, or a wide range of different grades. Even though this would also be my favorite in an ideal world, I regret how range voting depends on differences in the meaning of grades among people, how it is manipulable – which causes strong inequalities among the voting power of different citizens according to their ability to manipulate. This argument, I admit, may be questionable in a reduced city council, but ruling out range voting seemed cautious in the absence of information on its size and composition.

I have eventually given just one approval in the vote on voting rules: one to approval voting. It is because I had the ability of approving other rules that my choice of giving just one vote was truly meaningful.

A.2 D. Felsenthal

I adhere to the Condorcet Principle as a normative principle when one must elect one out of three or more candidates. This principle prescribes that should a candidate defeat every other candidate in pairwise comparisons (a Condorcet winner), it must be elected, and should a candidate be defeated by every other candidate in pairwise comparisons (a Condorcet loser), it must not be elected. This principle conveys the fundamental idea that the opinion of the majority should prevail, at least when majority comparisons pinpoint an unambiguous winner and/or an unambiguous loser. The Condorcet Principle takes into account only the ordinal preferences of every voter between any pair of alternatives because attempting to take into account also voters' cardinal preferences (as under the Range Voting procedure) would not only imply that a Condorcet winner may not be elected or, worse, that a Condorcet loser may be elected, but also that inter-personal comparisons of utility are possible and acceptable – which they are not!

I rank all the competing procedures for electing one out of m candidates ($m \geq 2$) according to two criteria: First,I prefer all Condorcet-consistent procedures over all procedures that are not Condorcet-consistent. Second, among Condorcet-consistent procedures I prefer those which are not vulnerable to non-monotonicity or to electing a Pareto-dominated candidate when a Condorcet winner does not exist; and among the procedures which are not Condorcet-consistent I prefer those which are not susceptible to one or more of the following four pathologies which I consider as especially serious: non-monotonicity, not electing a candidate who constitutes the top preference of an absolute majority of the voters (aka absolute Condorcet winner), electing a candidate who is a Condorcet loser or is Pareto-dominated.

According to these criteria I approved only Kemeny's and Copeland's procedures because they are both Condorcet-consistent and are not susceptible to any of the above mentioned four pathologies.

My rank-order of the 18 competing procedures is as follows:

Kemeny > Copeland > Black > Nanson > Untrapped Set > Fishburn > Uncovered Set > Top Cycle > Simpson > Borda > Coombs > Alternative Vote > 2-round Majority > Plurality > Majority Judgment > Approval Voting > Leximin > Range Voting.

A.3 W.V. Gehrlein

In all honesty, I do not remember exactly which of the many possible rules that were listed that I voted for during this impromptu exercise. However, my general convictions were expressed on the ballot that I submitted. The first statement on my ballot was: "In a perfect world I would recommend any Condorcet consistent voting rule". Standard arguments against the implementation of majority rule based voting are too heavily focused on one atypical example of something that could conceivably happen to ignore an "almost-majority" minority voting bloc with strong preferences. The obvious question is: What is the likelihood that such a scenario would ever actually exist? We all know that such hypothetical voting situations can always be developed to make any voting rule appear to behave very poorly on some criterion. The only practical way out of this dilemma must therefore be based on the likelihoods that voting rules display such bad behavior. In the context of evaluating voting rules to elect the mayor of a city in a typical situation, my assumption from scenarios that I am familiar with would make the possibility negligible that there would ever be more than four candidates. Since there is a very high probability that a Condorcet winner will exist in such cases, why should we not elect that candidate?

The answer to the immediately preceding question is that Condorcet consistent procedures are not always easy to implement with a larger number of candidates, which led to my second statement on the ballot. “In the real world I would recommend (some elimination rules that I do not recall) and Borda Rule”. These rules would give a reasonable probability of electing the Condorcet winner, while also being both explainable to and acceptable to the electorate, without any implication that simplicity should be the only criterion for evaluating voting rules. Arguments about the relatively large probability with which some of these voting rules can be manipulated are typically based on the assumption that one group of voters with similar preferences can manipulate the outcome, while all other voters are completely naive to the situation. When it is further assumed that these other voters are aware of such possibilities and that they can react accordingly, the probability that the winner could actually be changed is significantly reduced. However, it is definitely reasonable to conclude from this exercise that plurality rule is not considered to be acceptable and that Approval Voting is the clear winner when voting is done by Approval Voting. But, it is critical that we must not forget the significant concerns that have been raised about the type of winners that are selected when Approval Voting is employed.

A.4 J.-F. Laslier

I do not adhere to the Condorcet principle as a *normative* principle; if 49% of the population strongly prefer A to B and 51% slightly prefer B to A, I think that A is collectively preferable. My first best decision rule is thus utilitarianism, or “range voting”. But I found Approval Voting a very good practical mechanism to approximately achieve the utilitarian outcome. For the practice, I find that Condorcet-consistent procedures advisable, except in the extreme but important case of a society split in two. The best Condorcet procedure to me is the randomized procedure studied by B. Dutta, G. Laffond, M. LeBreton and myself under the name Essential set, but this rule was not proposed. In most cases, the Simpson rule (Minmax procedure) is a good way to select in the Essential set, like Kemeny, Coombs, and others. My preference was:

Range > Approval > various Condorcet methods among which I make little difference > Two round plurality > Alternative vote > Leximin > Majority Judgment > Plurality.

My guess was that, for this election, Approval would win, maybe challenged by Alternative vote (I was right!). Therefore I voted for Approval and Range. Here is my complete ranking, with my sincere utilitarian view scaled on the 0–100 scale:

Range (100) > Approval (99) > Kramer-Simpson (85) > Coombs (84) > Kemeny (83) > Copeland (82) > Nanson (81) > Black (80) > Borda (50) > Fishburn (21) > UncoveredSet (20) > 2-roundMajority (18) > AlternativeVote (17) > UntrappedSet (16) > TopCycle (15) > Leximin (10) > Majority Judgment (1) > Plurality (0)

A.5 M. Machover

I consider that decision about which voting procedure should be used must be governed by some meta-principle. I also consider that an appropriate meta-principle for the present hypothetical case is majority rule. I therefore gave my approval only to Condorcet-consistent procedures, selecting those that have additional desirable properties: Copeland's and Kemeny's procedures.

A.6 V. Merlin

While considering the question "what is the best voting rule that the city council of your town should use to elect a mayor?" my first reaction is that the procedure should be simple and easily understandable by the whole population of the city. The second question to answer is to which degree the Condorcet principle should be implemented. I do not adhere to the Condorcet principle, as a majority of 50% plus epsilon can impose a candidate which is the worst choice of the other voters, without considering compromise candidates. But at least, I do consider that a Condorcet loser should never be elected. Hence, Plurality rule is the worst system in the list.

So, I decided to advise Plurality with two rounds, Alternative Voting and Approval Voting. As long as there is a final duel, any elimination system using the plurality tallies will never elect the Condorcet loser. Plurality with two rounds and Alternative Voting are such systems. They are easy to explain, and have been implemented in different countries (France, Australia), with no major complaints. Moreover, Alternative Voting is hardly manipulable. I also consider that $k+1$ rounds before the final duel are better than k! Though I also voted for Approval Voting, it may be possible for it to select a Condorcet loser, if everybody just reports his first choice. But I think that the risk is quite limited, provided that a sufficiently large part of the population votes sincerely. Experiences show that voters tend also to approve more than one candidate. What would make me rank Approval Voting slightly below the two previous rules, is the fact that it has not been widely used in political elections. I felt that we still need more real life experiences to check that everything goes right with approval voting, but I am ready to give it its chance.

The simplicity argument goes against many Condorcet-consistent rules. Though Kemeny is an extremely elegant solution to the voting problem, it is rather sophisticated. For those who think that the Condorcet criterion should be implemented, I would recommend the Copeland method, which could be easily explained to the voters, as a tournament among the candidates.

At last, I fear that rules like the Borda count or Range voting could lead to undesired outcomes, when a fraction of the voters tries to manipulate it.

A.7 N. Miller

I cast approval votes for Approval Voting and Copeland. My votes did not reflect any general normative principle but rather my sense as to what would be both practical and reasonable for the type of election that Dan Felsenthal stipulated, namely the election of a mayor when a number of candidates are the ballot. A year ago I might have approved of AV/IRV also, but I now think that its problems are quite serious (even in practice, not just in theory). And, as a practical matter, my highest preference would be for Approval Voting, because it is simple to explain to voters, simple to cast votes, and simple to count. Moreover, most voters (in the US at least) would want to see some kind of vote totals in the newspaper the next day, which Copeland does not provide.

While Plurality lost our vote by a landslide, it works perfectly well in most US partisan general elections, since Duverger's Law works so powerfully that there are, literally or effectively, only two candidates in most such elections (the recent Senate contests in Florida and Alaska being notable exceptions). However, Approval Voting might be a definite improvement over Plurality in party primary elections and non-partisan general elections (which is how many mayors are elected), where often three or more candidates are on the ballot.

Finally, voting procedures need to be evaluated not only in terms of their "static" social choice properties (e.g., Condorcet consistency, monotonicity, etc.) but also in terms of their "dynamic" effects, e.g., incentives for candidate entry, candidate ideological positioning, etc., which affect the types of preferences profiles that are most likely to arise.

A.8 H. Nurmi

We were asked to propose voting systems that we could recommend or approve of to be adopted in the mayoral elections of our municipality. Recommend and approve of are two different – albeit related – things, but since we were asked to submit approval ballots, I felt encouraged to suggest more than one system (which I would NOT do if I were asked to recommend "a system"). I proposed Borda, Nanson and probably also Kemeny (someone may have preempted me on the latter, though). Anyway, my ranking is Nanson > Kemeny > Borda > approval voting and these (as far as I now recall) were on my ballot. Nanson and Kemeny are both pretty resistant to misrepresentation of preferences and take into account a great deal of the preference information given by the voters. (One could also point out that they are Condorcet, but I'm not much moved by that property any longer: some systems are vulnerable to adding or removing or cloning alternatives (e.g., Borda) (as shown by Fishburn), others to adding or removing voters with completely tied preferences (Condorcet) (as shown by Saari). Overall, being based on strict majority principle is not a decisive feature in my book. Although it can be argued that it is preferable to be ruled by a majority than by a minority, I think one should also sail clear of the dictatorship of majority. They (Nanson and Kemeny) both do well in terms of several choice theoretic criteria. Borda's advantage is in intuitively plausible metric rationalizability: it looks for the closest (in terms of inversion metric) consensus profile (in terms of the first ranked alternative) and since we are looking for a single winner, this makes sense. Borda count also does well in minority protection (as shown by Nitzan). Approval voting was also on my ballot, not so much because of its choice-theoretic properties, but because of intuitive appeal of its results: it sounds nice to have a mayor who is deemed acceptable by more voters than any other. I must say, though, that the interpretation of "approvability" is not obvious (and this pertains to the interpretation of our balloting result as well). Does the fact that I approve of a candidate mean that I can tolerate him as the mayor without resorting to active resistance or does it mean that I positively support him/her? I think this is what makes the approval voting results hard to interpret, but I guess a mayor that is even tolerated by more voters than any other candidate has at least tolerable prospects.

A.9 F. Plassmann

I view voting as a useful mechanism for making collective decisions when unanimous agreement is not possible. Elections should generally be preceded by discussions about the candidates and the importance that the voters attach to the election. If a minority of voters feels strongly about some candidates while the other voters are almost indifferent between these candidates, then it should be possible for the minority to convince sufficiently many of the others to change their minds prior to the vote-casting process. (I believe that in cases of near-indifference, most people's desire to preserve social harmony trumps rent-seeking.) If it is not possible to change sufficiently many voters' minds, then I would interpret this as evidence that the intensity in preferences between the groups is not as disparate as it might appear. I therefore feel comfortable ignoring voting rules that take account of the intensities of voters' preferences.

I value the Condorcet principle, and I see the main issue as what we should do when there is no Condorcet winner. Apart from the fact that it is not Condorcet consistent, the Borda rule has many attractive properties. Thus my first choice is Black's rule, which seems to be least susceptible, among many popular voting rules, to a wide range of voting paradoxes and which has a very small frequency of ties (as preliminary research with Nic Tideman suggests). The discontinuity of Black's rule also makes strategizing difficult. However, the need to understand two separate evaluation criteria might make Black's rule too complicated for some voters. Voters will accept the outcome of an election only if they understand how the ballots are to be counted. Approval voting is very simple and avoids some of the most egregious shortcomings of the plurality rule. Thus I would endorse approval voting in situations when simplicity is important.

A.10 M. Salles

I voted for Approval Voting and for Borda. I share Jean-François' view regarding the difficulty concerning majority rule. However, I do not go as far as him and would not recommend "range voting". In case there are a sufficient number of candidates, the Borda rule proposes a way to deal somehow with intensity of preferences without going as far as "Range Voting". Also I think that the voting method must be simple enough to be understood by the quasi-totality of the voters, which might not be the case of the alternative vote system or Kemeny's rule.

A.11 N. Tideman

A group of experts on voting theory wanted to learn their collective judgments of a variety of voting rules. They decided (by something like acclimation) to proceed by using approval voting. I thought this was a reasonable way of learning the general level of support for different voting rules, as a prelude to future discussion. I would not have recommended approval voting as a way to make a collective judgment of which voting rule is best. That, I think, requires both more time and a procedure for ranking the options, so that direct paired comparisons can be made.

I am quite startled by the high level of support for approval voting as a way of electing a mayor. What I find particularly distressing about approval voting is that it requires a voter to decide whether to draw a line between generally acceptable and unacceptable candidates, or to leave that task to other voters and instead to draw a line between the very best and the close contenders who are not quite as good. I think that voters for a mayor should not be required to choose between drawing those two types of lines.

The relevant criteria for a voting rule for mayor, in my opinion, are:

- First, the capacity of the rule to gain the trust of voters. This depends on the reasonableness and understandability of the logic of the rule and the ease with which the counting process can be followed. Investigating this requires psychological methods as well as knowledge of the logic of voting procedures.
- Second, the likely statistical success of the rule in identifying the outcome with the greatest aggregate utility, under the assumption that voters vote sincerely. This is something that can be investigated by statistical methods.
- Third, the resistance of the rule to strategic voting. This too can be investigated by statistical methods.

It is my guess that the best rule, by some intuitive averaging of these criteria, is the Simpson rule. But the empirical work that would justify this guess remains to be done.

A.12 W. Zwicker

When I suggested we vote on voting rules and use Approval Voting, I thought the proposal could not pass – we'd surely split over the use of Approval Voting. At the time, however, our "rump session" discussion was stuck and it seemed that a conversational grenade might do more good than harm. I was very surprised that no one objected; some, as one might expect, were enthusiastic. Then I realized the exercise might be constructive if we could collectively endorse the principle that plurality rule was terrible... despite the stated goals of our workshop, I'd never thought it likely that we'd reach even a loose consensus on a single alternative. My own ballot approved a large number of rules, for two reasons: I doubt that the current state-of-the-art allows us confidently to select a small number of best rules, and my genuine indecisiveness was consistent with the best strategy for making plurality look bad. In terms of my specific approvals, it seems like false comfort to rely on any single absolute principle as a guide, when every choice of a voting rule entails trade-offs along many dimensions, about which our understanding is limited. For example, I feel the draw of Condorcet's principle but reject it as an absolute, in part because some recent results suggest trade-offs between that principle and any reasonable degree of decisiveness. I've come to view decisiveness as an under-valued trait – very important, though not decisively so of course. I did approve some Condorcet extensions, but not top-cycle, because of its striking indecisiveness. Mathematically, Kemeny is beautiful whereas Black is plug-ugly, but I swallowed hard, approved Black, and disapproved Kemeny (because Kemeny winner are rankings, not individual candidates, and I can imagine what would happen the first time some real world election yielded a tie among several rankings).

Appendix B: 18 Voting Rules

In what follows, the "majority tournament" is the binary relation among candidates: "More than half of the voters prefer *a* to *b*". In that case we say that *a beats b* (according to pair-wise majority rule).

B.1 Approval Voting [App]

Each voter approves as many candidates as she wishes. The candidate with the most approval is elected. See Brams and Fishburn (1983), Laslier and Sanver (2010).

B.2 Alternative Vote [Alt]

Each voter submits a ranking (possibly incomplete) of the candidates. One first counts the number of times each candidate appears as top-ranked (his plurality score). The candidate with the lowest plurality score is eliminated. In a second count, the votes for this candidate are transferred to the second-ranked candidate (if any) on these ballots. The process is then repeated again and again until one candidate is ranked first by an absolute majority of the votes (original or transferred) is elected. See Farrell (2001), Farrell and McAllister (2006). Other names for this procedure or its variants: "Hare" system, "Single Transferable Vote", "Instant runoff".

B.3 Copeland [Cop]

Each voter submits a ranking of the candidates. For each candidate one computes his pairwise comparison score, that is the number of challengers this candidate beats under pair-wise majority rule. The candidates with the largest score are chosen. This Condorcet-consistent rule does not specify how ties (which are common when there is no Condorcet winner) are broken. See Laslier (1997). Other name: Tournament score.

B.4 Kemeny [Kem]

Each voter submits a ranking of the candidates. The rule defines a summary ranking as follows. For any ranking R of candidates one computes the sum, over all pairs (a, b) of candidates of the number of voters who agree with how R ranks a and b. Then R^* is chosen to maximize this total number of agreements. The elected candidate is the top-ranked candidate according to R^*. This procedure is Condorcet-consistent. See Young and Levenglick (1978), Young (1988). Other name: Median ranking.

B.5 Two-Round Majority [2R]

Each voter votes for one candidate. If a candidate obtains an absolute majority, he is elected. If not, a runoff election takes place among the two candidates who obtained the most votes. This rule is the most common rule throughout the world for direct elections, but it has seldom retained the attention of social choice theorists. See Lijphart (1994), Blais et al. (1997,2010), Taagera (2007). Other name: Plurality with a run-off.

B.6 Coombs [Coo]

Similar to the Alternative Vote but, at each round, if no candidate is ranked first by an absolute majority of the ballots, the eliminated candidate is the one who is most often ranked last. This procedure is Condorcet-consistent in the single peaked domain. Coombs (1964).[7]

B.7 Majority Judgement [Bal]

Each voter grades each candidate according to some pre-specified finite grading scale expressed in verbal terms. For each candidate one computes his median grade. Among the candidates with the highest median grade, a linear approximation scheme (described in Balinski and Laraki 2007) is used in order to choose the elected candidate. See Basset and Persky (1999), Gerlein and Lepelley (2003), Felsenthal and Machover (2008), Laslier (2011). Other names for this procedure or its variants: “Robust voting”, “Best median”.

B.8 Simpson [Sim]

Each voter submits a ranking of the candidates. The pair-wise vote matrix is computed. Then the chosen candidate is the one against which the smallest majority (in favor of another candidate) can be gathered. See Simpson (1969). Other names: “Minimax procedure”, “Simpson-Kramer rule”.

B.9 Borda [Bor]

Each voter submits a ranking of the candidates. For K candidates, each one receives $K-1$ points each time he is ranked first, $K-2$ points each time he is ranked second, etc. The elected candidate is the one who receives the largest number of points.

B.10 Black [Bla]

Choose the Condorcet winner if it exists and the Borda winner if not. Suggested by Black (1958).

B.11 Nanson [Nan]

Each voter submits a ranking of the candidates. The Borda score is computed. Candidates with Borda score equal to or below the average are eliminated. Then

[7] Thanks to Dan Felsenthal for pointing to me details of this definition.

a new Borda count is computed, on the reduced profile and the process is iterated. This procedure is Condorcet-consistent. See Nanson (1883).

B.12 Range Voting [RV]

Each voter gives to each candidate as many points as she wishes between zero and, say, 10 points. The elected candidate is the one who receives the largest number of points. Range Voting is not often considered in the voting rule literature since, from the theoretical point of view, it is essentially plain utilitarianism. See Arrow et al. (2002), Dhillon and Mertens (1999), Baujard and Igersheim (2010) or the `rangevoting.org` web site. Other names for this procedure or its variants: "Utilitarianism", "Point voting" and in French: "vote par note", which just means "voting by grading."

B.13 Top Cycle [TC]

Each voter submits a ranking of the candidates. The majority tournament is computed. The Top-Cycle is the smallest set of candidates such that all candidates in this set beat all candidates outside this set. This Condorcet-consistent rule does not specify how ties (which occur when there is no Condorcet winner) are broken. See Schwartz (1972), Laslier (1997).

B.14 Uncovered Set [UC]

Each voter submits a ranking of the candidates. The majority tournament is computed. A candidate a belongs to the Uncovered set if and only if, for any other candidate b, either a beats b or a beats some c who beats b. This Condorcet-consistent rule does not specify how ties (which occur when there is no Condorcet winner) are broken. See Miller (1980), McKelvey (1986), Laslier (1997). Other name (in the graph-theory literature): "Kings procedure".

B.15 Leximin [Lex]

Each voter grades each candidate according to some pre-specified grading scale. Each candidate k is evaluated according to the worst grade he received, say $g(k) = \min_v g(k, v)$. The elected candidate is the one with the best evaluation $g^* = \max_k g(k)$. If several candidates have the same evaluation g^*, the elected candidate is the one who receives g^* the least often. This rule is an important benchmark for normative economics. See Arrow et al. (2002).

B.16 Fishburn

This choice correspondence is a variant of the Uncovered set which is useful when the majority relation contains ties (exactly as many voters prefer a to b than b to a). See Aleskerov and Kurbanov (1999).

B.17 Untrapped Set

This choice correspondence defined by Duggan (2007) is a variant of the Top-Cycle which is useful when the majority relation contains ties (exactly as many voters prefer a to b than b to a).

B.18 Plurality

Each voter votes for one candidate. The candidate with the most votes is elected. This is the most common voting rule in the Anglo-saxon world and the literature is very large. Other name: First Past the Post.

Appendix C: Statistical Significance of the 3D Representation

The method for spatial representation of data sets is derived from multivariate factor analysis. Given is a symmetric matrix of positive numbers, intended to measure the distances between the items, say $dist(c, c')$. If each item c is represented by a point $\phi(c)$ in the Euclidean space of dimension d one can compute the sum of the squares of the distances between the items:

$$\sum_{c,c'} dist^2(c, c'),$$

called the total variance, and compare this sum to the sum of squares of the distances between the corresponding points:

$$\sum_{c,c'} (\phi(c) - \phi(c'))^2,$$

called the explained variance. The best representation with d dimensions can be computed numerically using linear algebra. The quality of the representation is measured by the ratio between explained and total variance. This technique was used for Approval Voting data by Laslier and Van der Straeten (2004) and by Laslier (2006).

Of course the quality of the representation can only increase with the number of dimensions. In the text I show a 3D representation that explains about 90% of the the variance. In order to check whether this figure should be considered as large, I replicated the same computation on randomly generated data. Recall that, with the real data the explained percentages are, respectively 39, 66 and 90 for 1, 2 and 3 dimensions.

In a first test, suppose that each voter approves of each candidate independently with a probability p that corresponds to the average approval rate (here: $p = 78/(15 * 22) \simeq 0.236$). Running 10.000 simulations I find that the observed figures $39\%, 66\%, 90\%$) are respectively attained with probability 0.016, 0.005 and 0.0005. It is thus clear that our data set has much more structure than a totally random one, in which all candidates are alike, up to random fluctuations.

In a second test, suppose that we set the expected number of approval votes received by each candidate c to its actual value. So suppose that each voter independently approves of each candidate c with a probability $p(c)$ equal to the actual approving percentage of this candidate. For instance for the candidate *Approval Voting*, $p(\texttt{App}) = 15/22 \simeq 0.6818$. We thus keep trace that some candidates are good and some are not, but we lose the correlation among candidates. In that case, I find that that the observed figures $(39\%, 66\%, 90\%)$ are respectively attained with probability 0.07, 0.07 and 0.03. Again one can conclude from this statistical test that it is not by chance that the real data set provides such large figures.

References

Aleskerov, F. & Kubanov, E. (1999). A degree of manipulability of known social choice procedures. In S. Aliprantis, A. Alkan, & N. Yannelis (Eds.), *Studies in economic theory* (Vol. 8), *Current trends in economics: Theory and applications* (pp. 13–28). Berlin: Springer.

Arrow, K. J. (1951). *Social choices and individual values*. New York: Wiley.

Arrow, K. J., Sen, A. K., & Suzumura, K. (2002). *Handbook of social choice and welfare* (Vol. 1). Amsterdam: Elsevier.

Balinski, M. & Laraki, R. (2007). A theory of measuring, electing, and ranking. *PNAS, 104*, 8720–8725.

Basset, G. W. & Persky, J. (1999). Robust voting. *Public Choice, 99*, 299–310.

Baujard, A. & Igersheim, H. (2010). Framed-field experiments on approval voting in the political context. In J.-F. Laslier & R. Sanver (Eds.), *Handbook on approval voting* (pp. 357–395). Heidelberg: Springer.

Black, D. (1958). *The theory of committees and elections*. Cambridge: Cambridge University Press.

Blais, A., Massicotte, L., & Dobrzynska, A. (1997). Direct presidential elections: A world summary. *Electoral Studies, 16*, 441–455.

Blais, A., Labbé-St-Vincent, S., Laslier, J.-F., Sauger, N., & Van der Straeten, K. (2010). Strategic vote choice in one round and two round elections. *Political Research Quarterly, 20*, 1–9.
Brams, S. J. & Fishburn, P. C. (1983). *Approval voting*. Boston: Birkhäuser.
Coombs, C. H. (1964). *Theory of data*. New York: Wiley.
Dhillon, A. & Mertens, J.-F. (1999). Relative utilitarianism. *Econometrica, 67*, 471–498.
Duggan, J. (2007). A systematic approach to the construction of non-empty choice sets. *Social Choice and Welfare, 28*, 491–506.
Dutta, B. & Laslier, J.-F. (1999). Comparison functions and choice correspondences. *Social Choice and Welfare, 16*, 513–532.
Farrell, D. M. (2001). *Electoral systems: A comparative introduction*. New York: Palgrave.
Farrell, D. M. & McAllister, I. (2006). *The australian electoral system: Origins, variations, and consequences*. Sydney: University of New South Wales Press.
Felsenthal, D. & Machover, M. (2008). The majority judgment voting procedure: A critical evaluation. *Homo Oeconomicus, 25*, 319–334.
Gehrlein, W. V. & Lepelley, D. (2003). On some limitations of the median voting rule. *Public Choice, 117*, 177–190.
Kilgour, M. K. (2010). Approval balloting for multi-winner elections. In J.-F. Laslier & R. Sanver (Eds.), *Handbook on approval voting* (pp. 357–395). Heidelberg: Springer.
Laffond, G., Laslier, J.-F., & LeBreton, M. (1993). The Bipartisan set of a tournament game. *Games and Economic Behavior, 5*, 182–201.
Laffond, G., Lainé, J., & Laslier, J.-F. (1996). Composition-consistent tournament solutions and social choice functions. *Social Choice and Welfare, 13*, 75–93.
Laslier, J.-F. (1997). *Tournament solutions and majority voting*. In *Series studies in economic theory* (Vol. 7). Heidelberg: Springer.
Laslier, J.-F. (2006). Spatial approval voting. *Political Analysis, 14*, 160–185.
Laslier, J.-F. (2011). On choosing the alternative with the best median evaluation. *Public Choice*, forthcoming.
Laslier, J.-F. & Sanver, R. M. (Eds.). (2010). *Handbook on approval voting*. Heildelberg: Springer.
Laslier, J.-F. & Van der Straeten, K. (2004). Election présidentielle: une expérience pour un autre mode de scrutin. *Revue Française de Science Politique, 54*, 99–130.
Lijphart, A. (1994). *Electoral systems and party systems: A study of twenty-seven democracies, 1945–1990*. Oxford: Oxford University Press.
McKelvey, R. D. (1986). Covering, dominance, and institution-free properties of social choice. *American Journal of Political Science, 30*, 283–314.
Miller, N. R. (1980). A new solution for tournaments and majority voting. *American Journal of Political Science, 24*, 68–96.
Nanson, E. J. (1883). Methods of elections. *Transactions and Proceedings of the Royal Society of Victoria, 19*, 197–240.
Schwartz, T. (1972). Rationality and the myth of maximum. *Noûs, 6*, 97–117.
Simpson, P. (1969). On defining areas of voter choice. *Quarterly Journal of Economics, 83*, 478–490.
Taagerera, R. (2007). Electoral systems. In C. B. Boix & S. Stokes (Eds.), *The oxford handbook of comparative politics*. Oxford, Oxford University Press.
Tideman, T. N. (1987). Independence of clones as a criterion for voting rules. *Social Choice and Welfare, 4*, 185–206.
Young, H. P. (1988). Condorcet's theory of voting. *American Political Science Review, 82*, 1231–1244.
Young, H. P. & Levenglick, A. (1978). A consistent extension of Condorcet's election principle. *SIAM Journal on Applied Mathematics, 35*, 285–300.