http://eprints.lse.ac.uk

The underlying assumptions of electoral systems

Moshé Machover

CPNSS, London School of Economics and Political Science

Revised January 2011

Please address all correspondence to:

Moshé Machover
5 Milman Road
Queen's Park
LONDON NW6 6EN
UK

Phone: +44 (0)20 8969 5356
E-mail: moshe.machover@kcl.ac.uk

Abstract

I propose a twofold classification of the main considerations underlying the choice
of an election procedure: political criteria on the one hand, and social-choice
criteria on the other. I formulate political dichotomies, each combination of which
narrows down the choice of procedure to a sub-class of the class of all procedures.
I discuss what social-choice theory has to offer in each of these.

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 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.

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) v 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 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.

---

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

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) v 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.

3 PR procedures

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

3.1 PR&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

---

[2]This goes back to John Stuart Mill. In [6, Ch. 7] he clearly advocates PR; but then seems to take it for granted that electing a legislature must use some form of DR.

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.

3.2 PR&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 [1].[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.

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,

---

[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 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 [6, Ch. 7]; cf. footnote 2.

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

[7]For a proof, see [5, Section 6.2].

[8]For a discussion of the technical properties and political advantages of this procedure, see [1]

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.)

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) v 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 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.

_____

and [5, Section 4.4].

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

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 (Ch. 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&DP) or allow lottery processing (that is, AR&LP).

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

The combination MR&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&DP.

This leaves the final combination:

4.1 MR&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, Laslier and Le Breton [4], and independently (using a quite different method) by Fisher and Ryan [3].

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.

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 [2], this provides an electoral procedure based purely on MR&LP: conduct a weighted lottery, in which each candidate x is assigned weight p_x.

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 combinations of these, then social choice can provide a single optimal procedure.

References

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

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

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

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

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

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