

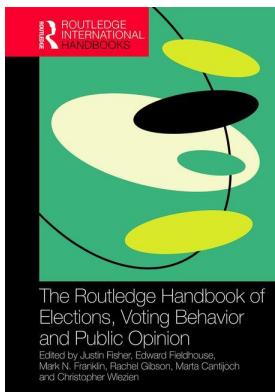
This article was downloaded by: 10.3.97.143

On: 03 Oct 2023

Access details: *subscription number*

Publisher: *Routledge*

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: 5 Howick Place, London SW1P 1WG, UK



The Routledge Handbook of Elections, Voting Behavior and Public Opinion

Justin Fisher, Edward Fieldhouse, Mark N. Franklin, Rachel Gibson, Marta Cantijoch, Christopher Wlezien

Electoral systems

Publication details

<https://www.routledgehandbooks.com/doi/10.4324/9781315712390.ch17>

Iain McLean

Published online on: 26 Sep 2017

How to cite :- Iain McLean. 26 Sep 2017, *Electoral systems from: The Routledge Handbook of Elections, Voting Behavior and Public Opinion* Routledge

Accessed on: 03 Oct 2023

<https://www.routledgehandbooks.com/doi/10.4324/9781315712390.ch17>

PLEASE SCROLL DOWN FOR DOCUMENT

Full terms and conditions of use: <https://www.routledgehandbooks.com/legal-notices/terms>

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

17

ELECTORAL SYSTEMS

Iain McLean

There is (usually) no best electoral system

In general, there is no best electoral system, but some are worse than others. Each electoral system appeals, openly or implicitly, to a concept of representation. There are two of these, explained in the second section of this chapter, namely the principal-agent concept, and the microcosmic concept. Both are valid, but they are incompatible with one another. Therefore, electoral system designers must choose a system appropriate for the context in which it is to be used. They must also decide whether the system is to aggregate judgments or to aggregate preferences, and whether it is to elect one person or a multi-member assembly.

This framework enables us to analyze the main varieties of electoral system. Within each family (e.g., majoritarian and proportional), some systems are no worse than their rivals, and are better in at least one respect.

Let us then start with some theorems, although we will not stay long with the mathematics of elections. The three theorems that every electoral system designer needs to know are:

- May's theorem (May 1952, 1953): simple majority rule is the only system that simultaneously satisfies some desirable conditions for choosing between two alternatives;
- The median voter theorem (Black 1948, 1958; popularized by Downs 1957): when voting opinion is one-dimensional, the position (person) preferred by the median voter will win in any well-designed voting system;
- Arrow's theorem (Arrow 1963 [1951]): no choice or aggregation system can simultaneously satisfy some minimal requirements of consistency and fairness.

May's theorem

An election rule is a mapping from many votes to a single decision, or choice. The first property we want it to satisfy is *decisiveness*. Imagine the rule as an obedient but cuddly robot charged with making the binary choice between X and Y. We want it always to be able to say "X won" or "Y won" or "X and Y have tied." No matter how numerous the individual votes that we have fed into it, we don't want it to keep on whirring and be unable to give us a result.

Next up is *anonymity*. A rule is anonymous if swapping the identity of two voters doesn't affect the result. Suppose I am one voter, and the King of Sweden is another. A rule is anonymous

if, after I have cast the King of Sweden's vote and he has cast mine, the result is unaffected. Sometimes we may want a rule to be non-anonymous – for example, when we give the chair of a meeting a casting vote to break a tie. But in a democracy we would normally want my vote to be worth neither more nor less than the King of Sweden's.

Neutrality is to options what anonymity is to voters. Suppose the initial outcome is X. Then let everyone who voted for X vote for Y, and everyone who voted for Y vote for X. With a neutral rule, if the voters flip, so should the outcome, which will now be Y. As with anonymity, we would sometimes want a non-neutral rule. A majority rule for juries is an example. Suppose the rule says that at least 10 out of 12 jurors are required to convict. Then if nine say Innocent and three say Guilty, the result is Innocent. If we flip, with nine saying Guilty and three saying Innocent, the result is still Innocent. So a jury rule may be non-neutral. But normally, in a democracy, we want our rules to be neutral.

Finally, *positive responsiveness*. A rule is positively responsive if, when the outcome is a tie, a switch of a single vote suffices to switch the overall result in the direction of the switch. Qualified-majority rules are not positively responsive. If the rule requires (say) a two-thirds majority for a motion to pass, and stipulate that in the event of a tie the motion does not pass, then a switch of one voter from against to for does not cause the motion to pass. Once again, there are sometimes good reasons for qualified-majority rules, but generally in a democracy we want the majority to win.

It is easy to see that simple majority rule satisfies the four conditions. What requires some math is to prove that it is the *only* rule that satisfies the four conditions. But it is (May 1952, 1953). That seems to constitute a powerful argument for majority rule.

The median voter theorem

Suppose that every voter can be ranked by her position on some scale. It might be left–right, but it could be anything (musical tastes, say). Let us assume that those who most like the music of J. S. Bach most hate the music of Justin Bieber, and vice versa. A Bach-lover likes a piece of music the less, the further down the Bach-to-Bieber scale it lies. A Belieber is the mirror image. A person whose favorite composer is somebody else will like music less, the further it departs from her favorite toward either Bach or Bieber. The technical terms for this condition are “un-dimensionality” and “single-peakedness” (Black 1958). When opinion is unidimensional and single-peaked, then with any good voting rule the favorite option of the median voter will beat all others. The median voter is the one with exactly as many voters to her “left” as to her “right.” In the world of easy proofs, the number of voters is always odd. Like May's theorem, the median voter theorem (MVT) makes a powerful argument in favor of democracy.

Arrow's theorem

Unfortunately, it is not always so simple. In a stunning result, Arrow (1963 [1951]) proves that no system satisfies some minimal requirements of fairness and logicity. These are: decisiveness (as in May); transitivity (if A is preferred to B and B is preferred to C, then A is preferred to C); the Pareto condition (if everyone prefers X to Y, the system does not rank Y above X); independence of irrelevant alternatives (the choice between X and Y is a function only of voters' preferences between X and Y, and nothing else); and non-dictatorship (there is no voter whose preference automatically becomes the output of the system, regardless of others' preferences).

This may all be a bit of a mouthful for the average electoral system designer, so I will try to spell out how it arises, and some of its implications. It has been known since Condorcet (1785,

1995 [1788]) that when there are at least three voters and at least three options, majority rule may be intransitive. I prefer A to B and B to C. You prefer B to C and C to A. She, over there, prefers C to A and A to B. By majority rule A beats B (I and She against You). B beats C (I and You against Her). Transitivity requires A to beat C. But C beats A (You and She against me). This may seem trivial, but it is not. It is called a “cycle” – a name conferred by Lewis Carroll (Dodgson 1876). It is the reason why the nice results of May and Black do not carry over to the multi-person, multi-option case with more than one dimension. Arrow’s theorem shows that there is no easy way out. Every system, tried, untried, or not yet invented, must violate at least one of Arrow’s conditions. Yet most of us would want at least those conditions, and much more besides.

It follows that there is no such thing as a perfect electoral system. The comforting thing about Arrow is that we can abandon the search now. For all we know, life may exist on another planet. But we know for sure that a system that meets all Arrow’s conditions does not exist, even on another planet.

That does not mean that any electoral system is as good as any other. It means that we have to decide, first, what an electoral system is *for*. Having decided that, we can say that some systems are better than others for that purpose.

Three sets of criteria for an electoral system

Two concepts of representation

The oldest meaning of the verb “to represent” in the authoritative *Oxford English Dictionary*, attested since 1390, is:

To assume or occupy the role or functions of (a person), typically in restricted, and usually formal situations; to be entitled to speak or act on behalf of (a person, group, organization, etc.); (in later use esp.) to act or serve as the spokesperson or advocate of.

(Oxford English Dictionary Online 2017)

I represent you if I stand for you in a place where you cannot: for example, if I am your lawyer, with expert knowledge that you lack; or if there are many of you, who cannot all attend parliament, so I attend it as your representative. This is now called the “principal-agent” conception. You are the principal(s), and I am your agent.

Almost as old (attested from 1400) is sense 8b: “To bring clearly and distinctly before the mind or imagination; to describe, evoke, conjure; to imagine, conceptualize” (*Oxford English Dictionary Online 2017*). This is the sense in which a painting or a play re-presents a character. When applied to a large body of people, it leads to the “microcosm” conception of representation. In the ferment of the French and American Revolutions, the Comte de Mirabeau and John Adams came (independently, I think) to the same idea: that in some sense the assembly should be a microcosm of the people who elected it (cited in McLean 1991: 173).

Etymologically, both senses are perfectly valid. But they are incompatible. One person cannot be a microcosm of many, as the many comprise different genders, different ages, different ethnicities, and have different interests and values. For legislative elections, the principal-agent conception points to single-member districts. The microcosm conception points to proportional representation (PR), and/or to quotas (requiring, for instance, that a minimum proportion of candidates have some gender or ethnic characteristic). Some PR systems, like the one used in Scotland, Wales, Germany, and New Zealand, combine a single-member and a proportional

component. In those systems, it is only the proportional component that can bring about microcosmic representation.

Judgments vs. preferences

The second way to ask what a system is for is this: *Is it designed to aggregate people's judgments, or people's preferences?* Consider a jury. In principle, all the jurors want the same thing, which is to find out the unknown truth: did the accused commit the crime as charged, or not? There should be no difference between conservative and socialist jurors, or between Bachians and Beliebers: each in principle wants the same thing.

A less pure example is selecting somebody for a job. All members of the selection committee want the best person for the job. But they may have different ideas as to the qualities of the ideal candidate, so it is not a pure case of judgment aggregation.

A mass election is quite different. Some voters are moved by interests; some by ideology; many by both. Interests and ideologies differ. There is no sense in which the voters can be said to be groping toward an unknown truth.

Electing one vs. electing many

The third dimension for judging systems asks: is this an election of one person, or of more than one? Obviously, the concept of proportional representation has no meaning if there is only one post to fill. A president cannot be male, female, black, white, rich, poor, straight, gay ... in the same proportions as the population. There is only one of him (of her).

If the task is to elect a multi-member body, such as a parliament, then the principal-agent and microcosmic conceptions point in different ways. As already noted, the former points to single-member districts, so that each group of principals knows for sure who their agent is. The latter points to proportional representation, which cannot be achieved in single-member districts for the reason just given.

The main system families

For aggregating judgments

When a judgment-aggregation task is truly binary, as in most jury systems (not in Scotland, where a third verdict labeled "not proven" exists), then the May and median voter theorems take us quickly to a result. Juries should use a form of majority rule. It should be anonymous, because each juror is assumed to be as well qualified as each other to discern the truth. Most jury rules, however, use a non-neutral procedure, in which more votes are required for a guilty verdict than for a not-guilty verdict. This is reasonable, on the basis that convicting an innocent person is regarded (by legal theorists) as worse than letting off a guilty person.

Most judgment-aggregation processes are not truly binary, however. When selecting somebody for a job or an honor, there are typically more than two candidates. Normally a rank-order is required, so that if the person at the top of the list declines an offer or fails a medical, the selectors can go to the next.

The modern foundations of the theory of elections were laid in the French Enlightenment, largely around the question of what electoral system academicians should use when voting in elections to the national academies. The debate was most intense in the *Académie royale des sciences*, of which the principal theorists Condorcet (1785) and Borda (1784) were both fellows.

They had radically different ideas of what sort of person should be elected (McLean, McMillan, and Monroe 1998: xxvii). Their radically different ideas hid behind a (superficially) polite disagreement about voting rules. The Condorcet rule makes pairwise comparisons among all pairs of candidates. A *Condorcet winner* is the candidate who beats every other in those comparisons. Unfortunately, because of the possibility of cycles, a Condorcet winner may not exist. Condorcet has a procedure to deal with this, but it is so complicated that it took two centuries to be understood (Young 1988). That rules out Condorcet procedures as a practical matter.

The Borda rule is well-known, although not necessarily under that name. It is used in the Eurovision Song Contest and some other sporting tournaments (also, it should properly be called the Cusanus rule, as it was proposed by Nicolas of Cusa in 1434 [McLean and Urken 1995: 77–78]). The Borda rule is a points rule. You award a fixed score (usually zero) to the candidate you like least, and a fixed interval (usually 1) for each candidate above that. Hence for n candidates, each voter scores them from $n - 1$ down to 0. Ties may be allowed, or may be forbidden. The Borda rule can be operated either way.

The Borda rule is very simple, but has an overwhelming defect, which is obvious to anyone who has watched the Eurovision Song Contest. If you want your favorite singer to prevail against her most dangerous rival, be sure to give the latter a score of 0. As it is common knowledge that sophisticated voters will do this, the Borda rule may quickly degenerate into a contest to see who is the smartest manipulator.

We know that something like this happened when the Borda rule was used to elect academicians to the *Académie royale des sciences*. When this defect was pointed out, Borda said (I assume plaintively), “My scheme is only intended for honest men” (Black 1958: 182). That the Borda rule is flagrantly manipulable arises from its violation of Arrow’s axiom of independence of irrelevant alternatives (IIA). Under Borda, one can raise, or scupper, the chances of A by changing one’s ranking of B and C. Alone of Arrow’s conditions, the meaning and importance of IIA is not immediately obvious, and some serious scholars (see, for example, Saari 1990) doubt whether it should be a requirement. I (and probably most voting theorists) would condemn them to a lifetime of watching reruns of the Eurovision Song Contest.

Thus neither the Condorcet nor the Borda rule can usually be recommended directly. The first fails the decisiveness test. The second is extremely manipulable. But both of the Borda and Condorcet *principles* are valuable. One of the problems (which is really Arrow restated) is that the principles are not always compatible. When the Borda rule and the Condorcet rule yield different answers, which is the “true” winner? That cannot be settled by appealing to either principle.

A new procedure proposed by Balinski and Laraki (2010) may help to break this deadlock. Their “Majority judgment” procedure is designed, as the name suggests, for judgment-aggregation procedures. In their many applications, they show how the rules for wine-judging and figure-skating tournaments have evolved as participants have learned by bitter experience not to use the simple Borda (ranking) rule.

Balinski and Laraki (2010) point out that in a judgment aggregation, what voters are asked to do is not directly to rank the wines, skaters, or politicians, but to grade them against some mental standard. It is also what school and university examiners are asked to do. In some exam subjects, such as mathematics, it is easy to score the papers. In others, such as history, it is not. In math, a mark of 70 percent has a direct interpretation (the candidate got 70 percent of the answers right). In history, a mark of 70 percent has no direct meaning. But, in both cases, the examiners are grading the candidate against some unspoken criterion (“this is the threshold above which students are likely to be able to do math at university,” or “this is the threshold for a good degree”).

Electing a politician can be framed as a judgment-aggregation task. In their best-known experiment, Balinski and Laraki (2010) conducted an exit poll among people who had just voted in a French presidential election. The actual electoral system is a two-stage runoff, to be discussed later. But Balinski and Laraki offered voters leaving the polling station a mock ballot, in which they were asked to grade each candidate by completing the sentence: “To be president of France I judge in conscience that this candidate would be” The grades were the exam grades used to report French school results, which were therefore familiar to all voters. The rubric’s mention of “in conscience” was designed to remind voters that they were not (only) being asked to choose the candidate whose politics they most liked but (also) the candidate most fitted to be president of France.

For judgment aggregation where there are only two possibilities (e.g., “Guilty” and “Not Guilty”), the theorems that we have reviewed give the answer. We should use majority rule: a non-neutral majority rule if we judged that the consequences of an error in one direction are worse than the consequences of an error in the other direction. Jury rules are binary. They may offer an exception to the generalization with which I started, that there is no best electoral system.

How big a majority should be required to validate an outcome? That is partly a matter of judgment, which an electoral systems expert cannot settle: how much worse is it that an innocent person should be convicted than that a guilty person should go free? But in other jury-like contexts, some further mathematics is available. Juries comprise people (or maybe computers or sensors) who make imperfect judgments about the true state of the world. If a sensor in a jumbo jet or a nuclear power station indicates a fault condition, there are two possibilities: that the plane (reactor) is faulty, and that the sensor is faulty. As with convicting the innocent, the possibilities are asymmetrical. A faulty sensor may lead the plane to make an unnecessary, unscheduled landing in South Dakota. A sensor that correctly warns of a fault, but is ignored, may lead to the plane crashing with the loss of all on board.

The Condorcet jury theorem (Condorcet 1785; Austen-Smith and Banks 1996; List and Goodin 2001) enables us to apply the theory of probability to situations like these. If we know the average reliability of a juror (sensor), then we can stipulate the required majority of observations above which we can be, for all practical purposes, certain that the true state of the world is that which the majority of sensors show.

When there are more than two options in a judgment-aggregation problem, the Condorcet jury theorem can be adapted to suit. Unfortunately, because of Arrow’s result, the May majority rule theorem cannot be. For tasks such as selecting a shortlist of candidates for appointment, or choosing a president of France (if these are considered as judgment aggregations), how if at all can we get around the Arrow restrictions?

We could use a Condorcet procedure, comparing each candidate with each other, and ranking the candidates in descending order of the number of victories they score. That has the drawback that there may be a cycle. This rules out a Condorcet procedure for something like a presidential election. It may be acceptable for selecting a shortlist for appointment – but if the top candidates remain in a cycle, it will have to be broken somehow.

We could, but should not, use a Borda procedure. Even for the task of electing academicians, Borda found that his rule was subverted by dishonest men.

The Condorcet rule is almost never used in practice, and the Borda rule tends to be used only in relatively frivolous contexts such as song competitions. Better than either are two rules that are not as well-known as they should be.

Approval voting (Brams and Fishburn 1983) is very simple. You vote for as many candidates as you approve of – anything from one to all minus one. To choose one person, elect the person

with the highest number of approvals. To elect more than one (e.g., a shortlist), do the same, going down the list until the predetermined number of places have been filled.

An attraction of approval voting is that it caters for the voter who has very strong views about who should (not) be chosen, and the voter who is content for a large number of candidates to be regarded as acceptable. Unlike Borda, it does not require the voter to make a strict ranking of the candidates. But that opportunity can be a curse as well as a blessing. Though it cannot wholly escape the Arrow trap, because no ranking procedure can, it beats direct use of either Condorcet or Borda. Unlike Condorcet, it always produces a result. It is less manipulable than Borda.

More radically, bodies such as shortlisting committees should consider the Balinski–Laraki procedure. Because it is a grading, not a ranking rule, it can plausibly claim to escape the Arrow trap. At least one national scientific academy uses a hybrid of the Balinski–Laraki and Borda rules for electing academicians (personal communication). This is the more striking, as the academy in question selected its rule before the Balinski–Laraki proposal was formally characterized. It seems to have lighted on it serendipitously.

None of these rules is much used in real-world procedures to elect a single person. These rules are discussed in the next section.

For electing one person: interest aggregation

It may be utopian to regard the voters in a presidential election as actually engaging in deciding “in conscience” who is best fitted to be the next president. No presidential election system uses either approval voting or the Balinski–Laraki rule. So let us start from the other end, and evaluate the rules that are actually used.

Some countries elect their president directly, others indirectly through an electoral college. In either case, proportional representation is irrelevant, for the reason already given, namely that the person elected cannot be male, female, black, white, straight, gay, Christian, or Muslim in the same proportions as the population. There is only one of him (her). The electoral rule for the election of the president of Ireland is described on a government website as follows: “The Presidential election is by secret ballot and based on proportional representation by the single transferable vote” (Government of Ireland 2016).

This is a category mistake. A presidential election cannot use proportional representation. The actual Irish system is what is known in the UK as alternative vote, in Australia as preferential voting, and in the USA as instant runoff. Under any of these names, each voter ranks the candidates. If one candidate wins more than 50 percent of the first places, s/he is elected. If not, candidates with the fewest first preferences are eliminated in succession and their supporters’ next available preferences, if any, are transferred until one candidate has more than half of the remaining valid votes.

A cousin of this method is the French runoff system. In the first round anybody may stand. If nobody gets more than half of the votes cast, a second round is held a few days later, to which only the top two vote-getters from the first round go forward. The winner at the second round is elected.

Alternative vote and runoff systems are very similar. The main difference is that in a runoff system voters have more information. Before the second round, they know how the first-round ballots were cast. More information is good in itself, but it also gives more incentives to strategic behavior.

This subtle difference pales into insignificance, compared to the enormous defect of both systems, which is as follows: *There is no guarantee that the person elected is either the Condorcet or the*

(*sincere*) Borda winner. As noted, the Condorcet and Borda principles are not wholly compatible, but nobody has suggested a credible rival criterion for determining the “true” majority winner when there are more than two candidates. To see why they violate both criteria, consider a candidate who is everyone’s second preference. In a unidimensional world where the median vote theorem (above) applies, it is quite likely that such a person will exist: a centrist whose own first preferences are meagre, but whom both Bachians and Beliebers, left- and right-wing voters . . . , prefer to the standard-bearer of the other side.

Such a candidate is likely to be the Borda winner because she scores, on average, highest on voters’ ballot papers. She is likely to be the Condorcet winner because in a tournament of pairwise comparisons, she would win them all. This is an implication of the median vote theorem. But under either alternative vote or a runoff system, such a candidate is likely to be eliminated after the first round. Balinski and Laraki (2010) show from survey data that this has very likely happened in recent French presidential elections.

Although elimination rules are bad, they are not the very worst of rules which purport to be democratic. They are better than plurality rule (“first-past-the-post”), under which first preferences for all candidates are counted, and the modal candidate is elected. The modal candidate is the most popular single candidate, whether or not she has won over 50 percent of the vote. For an election to a single post, this is even worse than an elimination rule. At least an elimination rule cannot choose a Condorcet or Borda loser (namely, a candidate who loses every pairwise contest, and a candidate with the lowest average score. In the elimination round the successful candidate must have beaten one other, and therefore cannot be a Condorcet or Borda loser). First-past-the-post can select a Condorcet or Borda loser. It is therefore highly problematic for elections to a single post. It has a role in elections to a legislature, discussed below.

Therefore, for direct elections of one person, the systems to consider are either approval voting or Balinski–Laraki. Neither is clearly better than the other, but both are better than elimination systems or first-past-the-post.

Several democracies elect their president indirectly. Germany and the USA are two examples. Under the German Basic Law of 1949, which was enacted under the tutelage of the post-war Allied powers, the president is elected by a Federal Convention, which meets for that sole purpose: “The Federal Convention shall consist of the Members of the Bundestag and an equal number of members elected by the parliaments of the Länder on the basis of proportional representation” (Federal Law Gazette 2014). This is a huge advance on the US Constitution, which provides for an Electoral College to elect the president. The Electoral College was one of the most contested items in the Federal Convention in Philadelphia in 1787. It was intended to assemble an intermediate body of wise men who would elect the president. It has never worked like that. American voters, when they vote in a presidential election, are technically voting for their state’s electors, not for the president. Since 1800, would-be electors have announced who they will support. This combines with the non-constitutional convention of winner-takes-all for Electoral College votes. In almost every state, the plurality-winning state wins that state’s entire Electoral College vote. This is one reason (there are many) why *election inversions* occur. An election inversion occurs when the winner of the presidential election is not the popular vote winner. The most momentous such inversion was in 1860, when Abraham Lincoln not only got less than 40 percent of the popular vote, but would have won in the Electoral College even in a straight fight with Stephen Douglas, who was the Condorcet and Borda winner among the four candidates. There were election inversions in 2000 and 2016. In 2016, Hillary Clinton (Democrat) got more popular votes than Donald Trump (Republican), who was elected. The US Electoral College has no friends among electoral system scholars (if that matters), but because part of

the arrangement is in the Constitution, it is unlikely to be revised (see further Peirce and Longley 1981; Miller 2012).

For electing many people: principal-agent systems

Finally, consider the case where the electoral system is aggregating interests, not judgments, and is used to choose a multi-member body such as a parliament. There are many handbooks to world electoral systems (see, for example, Colomer 2004). Here we have space only to concentrate on fundamentals. There cannot be a best electoral system, because of the incompatible conceptions of representation discussed above. The principal-agent conception requires single-member districts. The principals (voters) need to know who is their unique agent (legislator), if only so that they have the chance to “throw the rascals out” at the next election. When it is objected that single-member districts produce highly disproportionate results in all the countries which use them, including the USA, UK, France, Canada, and India, the principal-agent theorist replies that the objection is based on a category mistake – the system is not intended to be proportional.

Sometimes, a single-member system uses an elimination rule. This applies in Australia, where the House of Representatives is elected by preferential voting (alternative vote). This arose in a quite unprincipled way. In 1918 the governing Nationalists had split over support for World War I. They faced a resurgent Labor Party in a by-election. Between the election being called and taking place, they rushed through a change to a preferential system where (they hoped) National supporters would transfer their support among varieties of National candidates and keep Labor out (McLean 1996).

In 2011, the UK Coalition government held a referendum on changing the Westminster electoral system from plurality to alternative vote. This was a sort-of compromise between the Liberals’ wish for proportional representation and the Conservatives’ wish for no change. It was not much of a compromise and was heavily defeated. Minor variants of alternative vote (AV plus; supplementary vote) have been proposed or tried. Supplementary vote was introduced for London mayoral elections in the hope of keeping them a two-party game between the Conservative and Labour candidates. These variants are described by the Electoral Reform Society (2010). There is no need to discuss elimination systems in detail here, because they barely improve on plurality rule. They may produce an even more disproportional result; they can penalize a centrist party that might be the Borda or Condorcet winner in each seat; against that, they have the advantage that they cannot elect a Borda or Condorcet loser.

For electing many people: proportional systems

To implement a microcosmic conception of representation, a proportional (PR) electoral system is often needed. Each of these tries to ensure that, as closely as possible, the seat shares in the legislature match the vote shares in the country. They fall into three main classes: *list systems*; *additional member systems*; and the *single transferable vote* (STV) family.

List systems

In these, the voter votes for a single party. The parties are assigned as near as can be, after rounding, the correct number of seats each. List systems can be highly proportional. They face two objections: that control of who is elected lies more with the party than with the voter; and that some list systems use an incorrect rounding-off algorithm.

When the parties have complete control of the order of the candidates on their lists (“closed-list”), they in effect decide who will be elected and who will not, at least if they can make a reasonable guess at the number of seats they will win. The formula for this is:

$$Q_i = \left\lceil \frac{V_i}{M+1} \right\rceil$$

where Q_i denotes the quota of seats for party i , V_i denotes the votes it has received, M denotes the number of seats to be filled (“district magnitude”), and the one-ended brackets $\lceil \quad \rceil$ denote the *upper integer bound* of the expression inside them. This means $V_i/(M+1)$ rounded up to the next whole number, or, if it is already a whole number, rounded up by 1.

Exactly how proportional a list system is depends on the district magnitude M . The larger is M , the more proportional the system. Therefore, the Netherlands and Israel, which treat the whole country as one district and M as the size of the whole legislature, have the most proportional electoral systems in the world.

To give less control to the parties and more to the voters, some countries use variants of “open-list,” where voters may choose a candidate within a party. Such votes are assigned both to the party and to the candidate within the party, hence giving voters some control.

Most list systems use an incorrect algorithm for rounding off fractional entitlements to seats. These algorithms are biased, and they do not guarantee that the right number of representatives is selected without some further tweaks. They use either:

- a system invented in the 1870s by the Belgian voting theorist Victor D’Hondt, which is identical to the system proposed in 1790 by Thomas Jefferson for rounding off entitlements to state seats in the US House of Representatives, or
- a system of giving a seat for each quota and then awarding any remaining seats in descending order of the fractional part of the quotient for unsuccessful candidates (largest remainder systems, invented in the USA by Alexander Hamilton in 1790).

The D’Hondt formula favors large parties, sometimes giving them more seats than their quotas. It is popular with the political parties who introduce PR, since usually, by construction, they are large parties. Hamilton systems sound fair, but they are bedeviled by paradoxes of monotonicity (where a candidate becomes more popular and thereby reduces her chances of election).

The only fair and non-paradoxical rounding-off algorithm is the one proposed in 1910 by the French mathematician André Ste-Laguë. This system, homologous to the rule proposed by Daniel Webster in 1832 for the US House problem, is the only one that treats large parties and small parties equally. It is therefore the only one that is fair to both the large-party and the small-party voter. The fascinating homologies (i.e., identities) between the American and European rules were first discovered, and proven, by Balinski and Young (2001), which is the authoritative source for everything in this paragraph.

Single transferable vote

This system was devised by several Victorian electoral reformers, the most prominent being Thomas Hare. Whereas list systems concentrate first on parties, STV concentrates first on fairness to voters. Hare’s idea was that every group of voters who amounted to at least a quota was entitled to a representative of their choice. In STV, voters cast ranked ballots. First preferences are counted, and any candidate(s) who have met the natural quota Q defined above are elected.

Their surpluses above Q are redistributed to the next available candidate listed on each ballot. When nobody else can be elected by this method, the candidates with fewest first preferences are (successively) eliminated until the required number of candidates is elected. STV is used in Ireland (both North and South), in Malta, for Scottish local government elections, and for the Australian Senate (although there a rule change in the 1980s, such that most voters choose their party's ranking rather than choosing their own, means that STV has mutated to open-list). Like largest remainder, it is non-monotonic, although in the case of STV that is not a serious practical objection for mathematical reasons (Bartholdi, Tovey, and Trick 1989). For more on STV in practice see, for example, Bowler and Grofman (2000).

STV can go wrong when it is applied in a judgment aggregation. For instance it is widely used in the internal elections of the Church of England. Since medieval times, voting theorists have characterized church elections as attempts to find out the will of God (Colomer and McLean 1998). But if opinions as to the will of God are incompatible and deeply entrenched, campaigners focus rather on the natural quota Q . A faction can obtain as many seats as it has quotas. Rather than discouraging factionalism, in a case like this STV may actually encourage it.

Single *non*-transferable vote (SNTV) was discussed by Lewis Carroll in the 1880s (McLean, McMillan, and Monroe 1996; Cox 1991). Under SNTV the voter has only one vote in a multi-member district. If voters and parties behave with full information and reasonable calculations about one another, however, its effects are similar to those of STV.

Additional member systems

List systems typically have large M and are responsive; STV has small M and achieves some of the advantages of single-member district systems at the expense of proportionality. A compromise between the two is to use a mixed-member system (MMS), as in Germany, New Zealand, Scotland, Wales, and the London Assembly. The details vary between these jurisdictions but the principle is the same. Part of the legislature is elected in single-member districts. The disproportionality that this produces is countered by electing the rest of the house in a regional or national party list, on a compensating rule such that the proportionality of the whole legislature is determined by the party share of the list vote. Since single-member district rules exaggerate the lead of the winning party, it follows that most of the list seats go to other parties.

There is a lot to be said for the MMS system as it tries to deliver the best of both worlds: preserve a principal-agent link while achieving (at least some) proportionality. It cannot of course achieve as much proportionality as a pure list system. This may have consequences as in Scotland in 2011, when the Scottish National Party (SNP) gained a majority of seats on about 44 percent of the votes. But this leads me to end where I began. No electoral system is perfect, but if the system designer starts by deciding the purpose of the votes and only then chooses an electoral system, she is proceeding in the right order.

Conclusion

So where have we reached? We have shown that there is no one answer to the question “What is the best electoral system?” That is because the question is incomplete. We should only ever ask, “What is the best electoral system *for this purpose?*” after deciding the purpose of the election. Is it to find out the truth? Is it to elect an executive? Is it to elect a parliament? Is it to choose a list, in order, of the best wines or the best figure-skaters? Is it to decide which candidate(s) for a job fit the essential criteria for that job, so that some of them are appointable, and others are not?

Electoral systems have consequences. We have seen that, with a given underlying structure of preferences, different electoral systems will produce different outcomes. Only one example is needed to make the point. The most important election in US presidential history, that of 1860, was won by Abraham Lincoln on less than 40 percent of the vote. He won because of the spatial distribution of the vote, because of the first-past-the-post system, and because of the Electoral College. Under Condorcet, Borda, or most other systems, the election would have gone to his great rival, Stephen A. Douglas (Riker 1982). The Civil War might or might not have taken place (Douglas died not long after the election), but it would have had a different course and perhaps a different outcome.

Thus electoral systems have fundamental implications for the effects of public opinion on policy-making and other activities of government. The public opinion that elected Lincoln was the same public opinion that would have elected Douglas under other systems. The history of the United States reached a fork in 1860. Every national election held since then has been held in the shadow of that contest. Equally, Balinski's and Laraki's experiments show that a different president would have been elected in France in 2002 had the electoral system been their majority judgment system. There is a huge downstream literature on the electoral effects of proportional versus majoritarian systems, which is out of scope for this chapter. But nobody should ever doubt the importance of the choices among the systems we have been discussing.

References

- Arrow, K. J. (1963) [1951] *Social Choice and Individual Values*, 2nd Edition, New Haven: Yale University Press.
- Austen-Smith, D. and Banks, J. (1996) "Information Aggregation, Rationality, and the Condorcet Jury Theorem," *American Political Science Review*, vol. 90, no. 2, March, 34–45.
- Balinski, M. L. and Laraki, R. (2010) *Majority Judgment: Measuring, Ranking, and Electing*, Cambridge, MA: MIT Press.
- Balinski, M. L. and Young, H. P. (2001) *Fair Representation: Meeting the Ideal of One Man, One Vote*, 2nd Edition, Washington, DC: Brookings Institution Press.
- Bartholdi, J. J. III, Tovey, C. A., and Trick, M. A. (1989) "Voting Schemes for Which It Can Be Difficult to Tell Who Won the Election," *Social Choice and Welfare*, vol. 6, no. 2, April, 157–165.
- Black, D. (1948) "On the Rationale of Group Decision-Making," *Journal of Political Economy*, vol. 56, no. 1, February, 23–34.
- Black, D. (1958) *The Theory of Committees and Elections*, Cambridge: Cambridge University Press.
- Borda, J.-C. (1784) "Mémoire Sur les Élections au Scrutin," in McLean, I. and Urken, A. B. (1995) *Classics of Social Choice*, Ann Arbor, MI: University of Michigan Press: 83–90.
- Bowler, S. and Grofman, B. (eds.) (2000) *Elections in Australia, Ireland, and Malta Under the Single Transferable Vote*, Ann Arbor, MI: University of Michigan Press.
- Brams, S. J. and Fishburn, P. C. (1983) *Approval Voting*, 2nd Edition, New York: Springer.
- Colomer, J. M. (ed.) (2004) *Handbook of Electoral System Choice*, Basingstoke: Palgrave-Macmillan.
- Colomer, J. M. and McLean, I. (1998) "Electing Popes: Approval Balloting and Qualified-Majority Rule," *Journal of Interdisciplinary History*, vol. 29, no. 1, Summer, 1–22.
- Condorcet, M. J. A. N. (1785) *Essai sur L'application de L'analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix*, Paris: Imprimerie royale.
- Condorcet, M. J. A. N. (1995) [1788] "Essai sur les Assemblées Provinciales," in McLean, I. and Urken, A. B. (1995) *Classics of Social Choice*, Ann Arbor, MI: University of Michigan Press: 139–168.
- Cox, G. W. (1991) "SNTV and D'Hondt are 'Equivalent,'" *Electoral Studies*, vol. 10, no. 2, June, 118–132.
- Dodgson, C. L. (1876) "A Method of Taking Votes on More Than Two Issues," in Black, D. (1958) *The Theory of Committees and Elections*, Cambridge: Cambridge University Press: 224–234.
- Downs, A. (1957) *An Economic Theory of Democracy*, New York: Harper and Row.
- Electoral Reform Society (2010) "Majoritarian Electoral Systems," www.electoral-reform.org.uk/majoritarian-electoral-systems, [accessed October 20, 2015].

- Federal Law Gazette (2014) "Basic Law for the Federal Republic of Germany," www.gesetze-im-internet.de/englisch_gg/englisch_gg.html, [accessed October 10, 2015].
- Government of Ireland (2016) "Presidential Election in Ireland," www.citizensinformation.ie/en/government_in_ireland/elections_and_referenda/national_elections/presidential_election.html, [accessed February 7, 2017].
- List, C. and Goodin, R. E. (2001) "Epistemic Democracy: Generalizing the Condorcet Jury Theorem," *Journal of Political Philosophy*, vol. 9, no. 3, September, 277–306.
- May, K. O. (1952) "A Set of Independent Necessary and Sufficient Conditions for Simple Majority Decision," *Econometrica*, vol. 20, no. 4, October, 680–684.
- May, K. O. (1953) "A Note on Complete Independence of the Conditions for Simple Majority Decision," *Econometrica*, vol. 21, no. 1, January, 172–173.
- McLean, I. (1991) "Forms of Representation and Systems of Voting," in Held, D. (ed.) *Political Theory Today*, Cambridge: Polity Press: 172–196.
- McLean, I. (1996) "E. J. Nanson, Social Choice, and Electoral Reform," *Australian Journal of Political Science*, vol. 31, no. 3, November, 369–385.
- McLean, I. and Urken, A. B. (1995) *Classics of Social Choice*, Ann Arbor, MI: University of Michigan Press.
- McLean, I., McMillan, A., and Monroe, B. L. (1996) *A Mathematical Approach to Proportional Representation: Duncan Black on Lewis Carroll*, Dordrecht: Kluwer.
- McLean, I., McMillan, A., and Monroe, B. L. (1998) *The Theory of Committees and Elections by Duncan Black; and Committee Decisions with Complementary Valuation by Duncan Black and R. A. Newing*, 2nd Edition, Dordrecht: Kluwer.
- Miller, N. R. (2012) "Electoral Inversions by the US Electoral College," in Felsenthal, D. S. and Machover, M. (eds.) *Electoral Systems: Paradoxes, Assumptions, and Procedures*, Berlin: Springer: 93–109.
- Oxford English Dictionary Online (2017) "Represent," <https://en.oxforddictionaries.com/definition/represent>, [accessed October 20, 2015].
- Peirce, N. and Longley, L. D. (1981) *The People's President: The Electoral College in America and the Direct Vote Alternative*, New Haven, CT: Yale University Press.
- Riker, W. H. (1982) *Liberalism against Populism*, San Francisco: W. H. Freeman.
- Saari, D. (1990) "The Borda Dictionary," *Social Choice and Welfare*, vol. 7, no. 4, December, 279–317.
- Young, H. P. (1988) "Condorcet's Theory of Voting," *American Political Science Review*, vol. 82, no. 4, December, 1231–1244.