EPISTEMIC DEMOCRACY:
ASSAYING THE OPTIONS*

Working Paper 2000-W9
30 June 2000

Christian List
Social & Political Theory Program
Research School of Social Sciences
Australian National University

and

Nuffield College, Oxford

and

Robert E. Goodin
Philosophy and Social & Political Theory Programs
Research School of Social Sciences
Australian National University

* © 2000 Christian List and Robert E. Goodin. Address for correspondence: Robert E. Goodin, RSSS, ANU, Canberra ACT 0200, Australia. Email goodinb@coombs.anu.edu.au, clist@coombs.anu.edu.au. This article was written during List’s tenure as Harsanyi Fellow in the Social & Political Theory Program, Research School of Social Sciences, Australian National University. Although it is a joint product in all the ways that matter, the initial hunch and much of the text are Goodin’s, the mathematics List’s. We are grateful to Geoff Brennan, Dave Estlund, David Firth, Jerry Gaus, Bernie Grofman, Iain McLean, David Miller, Philip Pettit and John Quiggin for helpful discussions.
Epistemic Democracy: Assaying the Options

Classical debates, recently rejoined, rage over the question of whether we want our political outcomes to be right or whether we want them to be fair. Democracy can be (and has been) justified in either way, or both at once.

For epistemic democrats, the aim of democracy is to "track the truth." For them, democracy is more desirable than alternative forms of decision-making because, and insofar as, it does that. One democratic decision rule is more desirable than another according to that same standard, so far as epistemic democrats are concerned.

For procedural democrats, the aim of democracy is instead to embody certain procedural virtues. Procedural democrats are divided among themselves over what those virtues might be, as well as over which procedures best embody them. But all procedural democrats agree on the one central point that marks them off from epistemic democrats: for procedural democrats, there is no "independent truth of the matter" which outcomes ought track; instead, the goodness or rightness of an outcome is wholly constituted by the fact of its having emerged in some procedurally correct manner.

1 Estlund (1990; 1993; 1997; 1998) is the most assiduous contemporary advocate of that position, but he has illustrious predecessors. Rousseau recommended democracy on the grounds that it tracks truths about the "general will" and "common good" (Rousseau 1762, bk 4, ch. 2; Barry 1964, pp. 9-14; Grofman and Feld 1988; Coleman 1989, pp. 204-5; cf. Estlund, Waldron, Grofman and Feld 1989; Miller 1992, p. 56). Nineteenth-century utilitarians advocated democracy on the grounds it tracked truths about "the greatest good for the greatest number" (Mill 1823).

2 As characterized in Cohen's (1986, p. 34) article "An epistemic theory of democracy," which is primarily responsible for introducing the term into the literature, "An epistemic interpretation of voting has three main elements: (1) an independent standard of correct decisions — that is, an account of justice or of the common good that is independent of current consensus and the outcome of votes; (2) a cognitive account of voting — that is, the view that voting expresses beliefs about what the correct policies are according to the independent standard, not personal preferences for policies; and (3) an account of decision making as a process of the adjustment of beliefs, adjustments that are undertaken in part in light of the evidence about the correct answer that is provided by the beliefs of others."


4 Coleman and Ferejohn (1986, p. 7) define "the proceduralist approach to the justification of collective decision" as one which "identifies a set of ideals with which any
Sometimes there is no tension between epistemic and procedural democrats, with all strands of democratic theory pointing in the same direction. That is the case where there are only two options before us. Then epistemic democrats, appealing to Condorcet’s jury theorem, say the majority winner is most likely to be the correct outcome. Procedural democrats of virtually every stripe agree. They, too, hold that majority voting is the best social decision rule in the two-option case; but their appeal is to the procedural, rather than truth-tracking, merits of majority voting. Although the many different rules that different procedural democrats recommend (Condorcet pairwise comparisons, the Borda count, the Hare or Coombs systems, etc.) might point in different directions in many-option cases, in the merely two-option case they do not. There, all the favorite decision rules of practically all democrats, procedural or epistemic, coincide on the majority winner.

This happy coincidence is confined to the two-option case, however. Where there are three or more options on the table, recommendations of the collective decision-making procedure ought to comply... [A] process of collective decision making would be more or less justifiable depending on the extent to which it satisfies them.... Proceduralism holds that what justifies a decision-making procedure is a necessary property of the procedure — one entailed by the definition of the procedure alone' rather than deriving from any calculation of consequences of applying that procedure. For elaboration of the jury theorem, see Section II below. In particular, May’s (1952) theorem is often adduced here, which shows that in a two-option case majority rule is the unique social decision rule satisfying some arguably compelling minimal conditions (decisiveness, anonymity, neutrality and positive responsiveness).

For definitions of all these different rules, see the key to Table 2 below. For further discussion of these and other decision rules, together with analyses of the extent to which they select the same outcomes, see Levin and Nalebuff (1995) and Merrill (1984). For a discussion of their formal properties from a social-choice theoretic perspective, see Riker (1983, ch. 4).

Indeed, as Borda (1784/1995, pp. 88-9) shows, the majority winner will also be the Condorcet and Borda winners whenever \( m > (k-1)/k \), where \( m \) is the proportion of votes the majority winner receives and \( k \) is the number of options over which they are voting. Note therefore that super-majority rules can have the same effect, in many-option elections (a two-thirds requirement in a three-option contest et seq.). Thus, in even in the many-option case, there is convergence between the recommendations of those other procedurally-favored rules and rule by supermajorities of a requisite size.
different strands of democratic theory diverge. Much modern writing on both social choice and electoral reform is dedicated to exploring the merits of alternate ways of aggregating people's votes into an overall social decision.\textsuperscript{9} Heretofore, however, those disputes have been conducted almost purely as intramural arguments within the proceduralist camp. Different social decision rules display different procedural virtues, and it is on that basis that we are typically invited to choose among them.

There is an epistemic dimension to that choice as well, however. It is a mistake to suppose (as philosophers writing about epistemic democracy sometimes seem to do\textsuperscript{11}) that their epistemic case for democracy based on the Condorcet jury theorem collapses where there are more than two options on the table. Anathema though it may be to some procedural democrats, plurality voting is arguably the simplest and possibly the most frequently used voting rule in many-option cases. Here we prove that the Condorcet jury theorem can indeed be generalized from majority voting over two options to plurality voting over many options.

That proof merely shows that the plurality rule is an "epistemically eligible" decision rule, however — not necessarily that it is uniquely preferred, epistemically.\textsuperscript{12} In addition to the proof of the truth-seeking power of the plurality rule, offered here for the first time, there has already been established a proof of the truth-seeking powers (in a much richer

\textsuperscript{9} Thus, Borda (1784/1995, p. 83) begins his famous paper initiating these debates with the words, "There is a widespread feeling, which I have never heard disputed, that in a ballot vote, the plurality of votes always shows the will of the voters. That is, that the candidate who obtains this plurality is necessarily preferred by the voters to his opponents. But I shall demonstrate that this feeling, while correct when the election is between just two candidates, can lead to error in all other cases."

\textsuperscript{10} Surveyed, respectively, in Mueller (1989) and Dummett (1985; 1997).

\textsuperscript{11} Listing reasons for thinking "the Jury Theorem approach to the epistemic value of democratic procedures is less than trustworthy," Estlund (1997, p. 189) includes the bald assertion: "the Jury Theorem assumes there are only two alternatives" (our emphasis).

\textsuperscript{12} Especially, as we shall see, when there is more information available about voters’ preferences than just a single vote for each voter.
informational environment) of the Condorcet pairwise criteria and Borda count.13

Where there are more than two options, different social decision rules seem to be differentially reliable truth-trackers. Furthermore, some rules seem to perform better under certain circumstances than others. We offer some sample calculations to suggest the dimensions and directions of these differences. But the differences are not great, and even the much-maligned plurality rule performs epistemically almost as well as any of the others where the number of voters is at all large (even just over 50, say).

We take no side in these disputes between epistemic and procedural democrats or among contending factions of proceduralists. We express no view on how much weight ought be given epistemic power compared to procedural virtues in choosing a social decision rule. Neither do we express any view on which procedural virtues are the most important for a decision rule to display. Our aim in this article is merely to "calibrate the epistemic trade-off" that might be involved in opting for one social decision rule rather than some other.

Our principal conclusion is that those epistemic trade-offs are not great. So long as the number of voters is reasonably large, virtually any of the social decision rules which have been commonly employed or recommended on democratic-proceduralist grounds seem to perform reasonably well (and nearly as well as any other) on epistemic-democratic grounds. A crucial corollary is that (assuming there is any epistemic dimension to politics at all) there is thus a strong epistemic case to be made

13 The proof of Young (1988; 1995), building on Condorcet (1785) himself, works in a very different way to ours, through sequences of pairwise votes over more than two options. For why we find our method preferable, see footnote 30 below. The "informational environment" must be richer for Condorcet pairwise comparisons and the Borda count, because those require us to know each voter's complete preference orderings over all options, whereas plurality voting requires only that we know each voter's first-choice preference.
for using democratic decision rules, whatever particular form we might prefer on democratic-proceduralist grounds.

I. Varieties of Epistemic and Procedural Democracy

As background to all those formal and computational results, let us first indicate briefly some of the key differences between and within theories of epistemic and procedural democracy.

A. Epistemic Democracy

The hallmark of the epistemic approach, in all its forms, is its fundamental premise that there exists some procedure-independent fact of the matter as to what the best or right outcome is. The epistemic approach tells us that our social decision rules ought be chosen so as to track that truth.

Where there is some decision rule which always tracks that truth without error, that can be called an epistemically "perfect" decision rule. It is hard to think what such a social decision rule might be. Democratic procedures, we can safely assume, will almost certainly never be commended as fitting that bill.

At best, democracy can be recommended as an epistemically "imperfect" decision rule. The defining feature of epistemically imperfect decision rules is that they track the truth, but they do so imperfectly. Their outcomes are often right, without being always right.

---

14 Except perhaps theocratic conceptions of truths revealed through god’s chosen spokesperson: but even then only if god reveals the truth, the whole truth and only the truth; and even then only if the person receiving god’s revelations never mistakes god’s words.
Where there is no epistemically perfect decision rule available, advocates of the epistemic approach must choose the best truth-tracker among the array of epistemically imperfect decision rules actually available. It is not self-evident that democratic procedures of any sort will necessarily be recommended on those grounds.\(^{15}\) Still, the epistemic virtues of information-pooling — which is what democracy amounts to, from this perspective — are such that democracy might lay a surprisingly strong claim to being the best imperfect epistemic procedure available.\(^{16}\) We say more on this score in subsequent sections, below.

**B. Procedural Democracy**

The hallmark of procedural approaches in all their forms is the fundamental premise that there exists no procedure-independent fact of the matter as to what the best or right social outcome is. Rather, it is the application of the appropriate procedure which is itself constitutive of what the best or right outcome is.

Procedures which are themselves constitutive in this way of the best or right outcome can be called procedurally "perfect" social decision rules.\(^{17}\)

\(^{15}\) Some say that epistemic criteria naturally favor "epistocracy" (rule by the epistemologically privileged alone) over democracy (rule by all the people, whatever their epistemic credentials). But given the epistemic advantages of information-pooling, it is sometimes better to pool information from more sources, even if that means drawing on less reliable sources (Young 1995, p. 53 n. 2); and Grofman, Owen and Feld (1983, p. 275) suggest that where the number of voters is large, assigning optimal weights them on the basis of competence might not produce very much better results, epistemically, than just giving them all equal weight regardless of competence.

\(^{16}\) On models of information-pooling more generally, as well as the Condorcet jury theorem as an example of it, see Grofman and Owen (1986). Waldron (1999, ch. 5) even finds a version of this argument in Aristotle's *Politics*, bk 3, ch. 11.

\(^{17}\) Or example, on the Austinian theory of sovereignty, for example, the "command of the Queen in Parliament" were fully constitutive of law and hence of political right, at least in that (limited) sense. Or, for another example, blunt democrats sometimes say the same: whatever comes out of a vote of Parliament, or a vote of the People, is the right outcome; there can be no other standard of what is politically best or right, beyond that.
Procedures which are not themselves constitutive of the right or best outcome, but which merely track (more or less perfectly) the outputs of those other (ideal-typical) procedures which are, can be called procedurally "imperfect" social decision rules.\(^{18}\)

Where the perfect procedure cannot itself be directly implemented as the social decision rule, advocates of the procedural approach must choose among the array of imperfect social decision rules that are actually available whatever decision rule best tracks the perfect procedure’s right-making dictates.

What attributes of social procedures, if any, should be regarded as constitutive in this way of politically best or right outcomes is of course a highly contentious issue — and one which we do not here propose to resolve.\(^{19}\)

Let us just offer an illustrative list of a few of the very different sorts of procedural criteria which have been recommended from time to time:

- Democratic proceduralists of an older and less formal bent have long insisted, among other things: that elections should be "free and fair," with voting proceeding without intimidation or corruption, and all valid ballots being counted; that the franchise should be broad, and elections regular and frequent; that the rules governing voting and elections should be common knowledge, and the procedure by which votes are transformed into decisions being publicly transparent (which is perhaps the main thing the "first past the post" plurality rule has going for it, procedurally\(^{20}\)).

\(^{18}\) For example, certain sorts of deliberative democrats say that the right thing to do is constituted by what would have been agreed in an “ideal speech situation,” which can only be imperfectly approximated by any actual political arrangements.

\(^{19}\) Any more than we attempted to resolve all the highly contentious philosophical issues in epistemology, in talking under the epistemological heading of "the truth."

\(^{20}\) “A voting method should be relative simple and transparent, both for voters and for those calculating the winner.... Simplicity helps explain why plurality voting is so widespread...” (Levin and Nalebuff 1995, p. 19).
Democratic proceduralists have also long insisted, with increasing precision and formality over the past couple of centuries, that the social decisions ought be systematically responsive to the preferences expressed by the voters; and democratic proceduralists have increasingly come to insist that that should be understood as meaning that they should be systematically responsive to "all the preferences of all the people" (which is what democratic proceduralists from Borda forward have had against plurality rule\textsuperscript{21}).

Democratic proceduralists of a more formal bent have recently added various axiomatic desiderata to the list of procedural criteria. They typically specify a set of (normative) minimal conditions that any acceptable social decision procedure should satisfy\textsuperscript{22}; and they then

\textsuperscript{21} "If a form of election is to be just, the voters must be able to rank \textit{each} candidate according to his merits, compared successively to the merits of \textit{each of the others}..." Borda does on to criticize plurality voting on precisely those grounds ("...the conventional form of election is highly unsatisfactory, because in this type of election, the voters cannot give a sufficiently complete account of their opinions of the candidates..."). Borda further recommends a system wherein "each voter ranks the candidates in order of merit" (or in which "we hold as many elections as there are combinations of candidates taken two by two so that each candidate can be compared to each of the others in turn" which as he says "is easy to see ... necessarily derives from the first") on the grounds that such a "method clearly gives us the \textit{most complete expression possible of the voters' opinions on all the candidates}" (Borda 1784/1995, p. 84, our emphases). As Dummett (1997, pp. 51-2) says, more simply, "Impartial reflection shows that the number of voters who think each candidate the worst ... is no less important ... than the number of voters who think each candidate the best" — and so on (we might add, in the spirit of his remarks) for every ranking in between (see more generally Dummett 1984, ch. 6; 1997, pp. 51-7).

Notice similarly the comments in Dahl's seminal \textit{Preface to Democratic Theory} (1956): "the principle of majority rule," Dahl writes, "prescribes that in chosing among alternatives, the alternative preferred by the greater number is selected. That is, given two or more alternatives $x, y, \ldots$, in order for $x$ to be government policy it is a necessary and sufficient condition that the number who prefer $x$ to any alternative is greater than the number who prefer any single alternative to $x$" (pp. 37-8). But instead of giving this Rule the plurality-rule interpretation toward which it naturally seems to tend, Dahl goes on to say, quite emphatically, that: "The \textit{essential requirement} of a system of voting that will satisfy the Rule is that voters ... must have an opportunity to vote for each alternative paired with another of the alternatives in a series of pairs sufficiently complete so that the alternative most preferred by a majority, if one such exists, will necessarily be selected. In some cases, this requires that a vote be cast on every pair of alternatives" (p. 43, our emphasis).

\textsuperscript{22} Such as Arrow's (1963) famous conditions of transitivity of social orderings, universal domain, the weak Pareto principle, independence of irrelevant of alternatives and non-dictatorship.
determine what social decision procedures, if any, satisfy these conditions.23

• Finally, democratic proceduralists might also include among their desiderata more pragmatic considerations of implementability, ranging from considerations of mere cost to ones of sheer computability.24 Some decision rules are informationally much more demanding than others, for example (the plurality rule requiring only information about people's first-choice preferences, whereas the others discussed below all require information about people's full preference orderings over all options).

Prioritizing different of those (and many other) procedural desiderata lead different procedural democrats to recommend different social decision rules. Since our concern in this paper is principally with the epistemic implications of that choice among social decision rules, we merely note those ongoing disputes within the democratic proceduralist camp and pass over them without further comment.

II. Extending the Jury Theorem: Plurality Voting over Many Options

23 As is well known, the set of social decision rules satisfying Arrow’s conditions (in a framework of ordinal preferences without interpersonal comparability) is empty. If universal domain is relaxed, the pairwise Condorcet criterion satisfies the other conditions, and if independence of irrelevant alternatives is relaxed, the Borda count satisfies the other conditions. For a more comprehensive overview of the main results of axiomatic social choice theory, see e.g. Sen (1982, esp. ch. 8). For application to voting rules specifically, see Riker (1983, ch. 4).

24 Condorcet (1792/1994, p. 218; 1792/1995, p. 145) acknowledges the impracticality of his exhaustive pairwise comparison procedure: "it is both awkward and time-consuming to form an initial judgment about the merits of the candidates and difficult to rank a large number of candidates in order of merit. Moreover, to extract these lists each voter’s opinion on all the candidates taken two by two and to use this to deduce a general result would be an immense and lengthy task.” Indeed, some voting procedures, including ones following from Condorcet’s and Lewis Carroll’s proposals, are not computationally feasible; see Bartholdi, Tovey and Trick 1989.
The Condorcet jury theorem, in its standard form, says this. If each member of a jury is more likely to be right than wrong, then the majority of the jury, too, is more likely to be right than wrong; and the probability that the right outcome is supported by a majority of the jury is a (swiftly) increasing function of the size of the jury, converging to 1 as the size of the jury tends to infinity.\textsuperscript{25}

Extrapolating from juries to electorates more generally, that result constitutes the jewel in the crown of epistemic democrats, who offer it as powerful evidence of the truth-tracking merits of majority rule.\textsuperscript{26} Much work has been done — by statisticians, economists, political scientists and others — to extend that result in many ways. It has been shown, for example, that a jury theorem still holds if not every member of the jury has exactly the same probability of choosing the correct outcome: all that is required is that the mean probability of being right across the jury be above one-half.\textsuperscript{27} It is also known, for another example, that a jury theorem still holds even if there are (certain sorts of) interdependencies between the judgments of different electors.\textsuperscript{28} The effects of strategic voting in a Condorcet jury context have also been studied extensively, showing mixed results.\textsuperscript{29} And so on.

What extensions and elaborations of the Condorcet jury theorem have almost invariably preserved, however, is the binary-choice form. (This is true in a way even of Peyton Young, who following Condorcet's own lead extends the theorem to cases of more than two options, but does so through a series of \textit{pairwise} votes between them.\textsuperscript{30}) The choice is thus typically

\textsuperscript{30}Like Condorcet (1785) before him, Young (1988; 1995) assumes that each voter has the same probability of making the correct choice in each pairwise comparison as each
between two options, or a series of options taken two-at-a-time. And in choosing between each of those pairs the average competence of voters is required to be over one-half.\textsuperscript{31}

Democratic theorists rightly remark that those constitute real limits on any epistemic case for democracy built on these foundations. As Estlund says, there is no reason to think that most important decisions in a democracy are going to boil down to two options (or, we might add, can be innocuously decomposed into a series of such two-option decisions).\textsuperscript{32} As Gaus says, there is no reason to think that people are generally more than half-likely to be right (particularly, we might add, where there are more than two options) — and the standard Condorcet jury theorem result works equally dramatically in reverse where they are more likely to be wrong than right, the probability of the collective choice being wrong growing exponentially with the size of the electorate.\textsuperscript{33}

Here we show that the Condorcet jury theorem can actually provide more comfort to epistemic democrats than previously imagined. Contrary to what they conventionally suppose, it can be extended even to plurality voting among many options. And contrary to what epistemic democrats

\textsuperscript{31} Or almost always so: as Grofman and Feld (1983, p. 271) show, the majority might be more likely to be right than wrong even if the average competence of voters drops to \( p_{\text{mean}}>0.471 \), although the value goes that low only when there are merely three voters.

\textsuperscript{32} Estlund 1997, p. 189. Shapley and Grofman (1984, p. 337) write, “The theorems discussed in this essay concern dichotomous choice, but this restriction may not be as serious as it might at first seem. If a group must choose from a set of alternatives (k\geq2), then it may do so by using any one of a number of binary choice procedures that decompose into sequences of pairwise (right fork or left fork) choices.” But as Riker (1983, p. 60) says, “Unfortunately, there is no fair way to ensure that there will be exactly two alternatives. Usually the political world offers many options, which, for simple majority decision, must be reduced to two. But usually ... the way the reduction occurs determines which two will be decided between. There are many methods to reduce the many to two; but, as has long been obvious to politicians, none of these methods is particularly fair ... because all methods can be rigged.”

\textsuperscript{33} Gaus 1997, p. 150.
conventionally suppose, voter competences can in those many-option cases drop well below fifty percent and the plurality winner still be most likely the correct choice.\(^{34}\)

We provide an extension of Condorcet’s jury theorem to the case of plurality voting over \(k\) options, where precisely one option (say, option \(i\)) is supposed to be the epistemically “correct” outcome. Specifically, we show the following:

(1) Suppose there are \(k\) options and that each voter/juror has independent probabilities \(p_1, p_2, \ldots, p_k\) of voting for options 1, 2, \ldots, \(k\), respectively, where the probability, \(p_i\), of voting for the “correct” outcome, \(i\), exceeds each of the probabilities, \(p_j\), of voting for any of the “wrong” outcomes, \(j \neq i\). Then the “correct” option is more likely than any other option to be the plurality winner.

(2) As the number of voters/jurors tends to infinity, the probability of the “correct” option being the plurality winner converges to 1.

The formal result is contained in Appendix 1. There are endless refinements and extensions of our result that might be made.\(^{35}\) And there are endless further issues that jury theorems, in all their forms, must eventually confront.\(^{36}\) For the purposes of this paper, we eschew these more technical issues, preferring to concentrate on the philosophical implications of

---

\(^{34}\) How can that be? After all, if the probability of each voter being right is less than half, is not the probability of each being wrong more than half? And in that case, does not the “reverse Condorcet” result set in, with a vengeance? The answer, of course, is simple. Under plurality rules the winner does not have to beat all the other options taken together. It has only to beat each of its rivals taken separately, where the opposition is divided \(k-1\) ways.

\(^{35}\) For example, we here assume that all voters have identical competence, that is, identical probabilities of voting correctly. The standard (majority voting, two option) Condorcet jury theorem has been generalized to cases of unequal levels of competence across voters/jurors; it is sufficient that their heterogeneous competences be symmetrically distributed around a mean which itself greater than 1/2 (Grofman, Owen and Feld 1983, pp. 268-9). Similar generalizations might be made in the \(k\)-option case.

\(^{36}\) Most especially, issues of strategic voting.
this extended Condorcet jury theorem in its simplest form for democratic theory more generally.

The major consequences of the result, as it bears on theories of epistemic democracy, would seem to be the following:

**Consequence 1:** The epistemically correct option is always more likely than any other option to be the plurality winner.\(^{37}\) The epistemically correct option may or may not be more likely than not to be the plurality winner, where there are more than two options on the table.\(^{38}\) But at least it is always more likely to be the plurality winner than is any other option.

**Consequence 2:** Where there are several options on the table, the plurality jury theorem can work even where each voter is substantially less than 1/2 likely to be correct, as required in Condorcet's original two-option formulation. The epistemically correct choice is the most probable among \(k\) options to be the plurality winner, just so long as each voter's probability of voting for the correct outcome exceeds each of that voter's probabilities of voting for any of the wrong outcomes. This implies that, if error is distributed perfectly equally, a better than \(1/k\) chance of being correct is sufficient for the epistemically correct option to be most likely to be the plurality winner among \(k\) options.

**Consequence 3:** So long as each voter is more likely to choose the correct outcome than any other, the correct option is more likely than any other option to be the plurality winner.

\(^{37}\) If all we know are people's first-preference votes, we can infer from Consequence 1 by Bayesian reasoning that the plurality winner is more likely than any other choice to be the epistemically correct option. Moving to a richer informational environment where we know people's full rankings over all options, as in the next section below, there will sometimes be choices (such as the Condorcet-pairwise or Borda winners) that are by similar Bayesian reasoning even more likely to be the epistemically correct options.

\(^{38}\) That depends on whether or not the number of voters/jurors is sufficiently large for the probability of the correct option being the plurality winner to be greater than 1/2.
regardless of how likely each voter is to choose any other option. Even if each voter is more than $1/k$ likely to choose each of several outcomes, the correct one is more likely to be the plurality winner than any other, just so long as the voter is more likely to vote for the correct outcome than that other outcome.

While the result says that the probability of the correct option being the plurality winner converges to certainty as the number of voters tends to infinity, it says nothing about how quickly that probability increases with increases in the size of the electorate. So far as epistemic democrats are concerned, how the function behaves at the limit — where the number of voters approaches infinity — is of less practical significance than how it behaves presented with plausible-sized electorates. The great boast of the Condorcet jury theorem in its traditional form is that the probability of the correct option being the majority winner grows quite quickly with increases in the size of the electorate. To what extent can the extended plurality jury theorem make the same boast?

To address that question, we present in Table 1 some illustrative calculations. (All these calculations are based on Proposition 1 in Appendix 1.) The first thing to note in Table 1 is this. Where each voter is more than half-likely to choose the correct option, the probability of the correct option being the plurality winner not only increases quickly with the size of the electorate: it increases all the more quickly in the $k$-option case than it does in the 2-option case. Where each voter is just over half-likely to choose the correct option in the $k=2$ case, the correct option has a probability of only 0.557 of being the plurality (majority) winner in an electorate of 51 voters; in the $k=3$ case, the probability of the correct option being the plurality winner in a same-sized electorate jumps to 0.937. In that (informal) sense, plurality winners in the $k$-
option case might be said to be even more reliable than majority winners in the 2-option case.

[Table 1 about here]
Table 1: Probability that the "correct" option (option 1) is the unique plurality winner

<table>
<thead>
<tr>
<th>number of options ($k$)</th>
<th>probabilities $p_1, p_2, \ldots, p_k$</th>
<th>probability that option 1 (the &quot;correct&quot; option) is the unique plurality winner for $n=$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>0.51, 0.49</td>
<td>0.527</td>
</tr>
<tr>
<td></td>
<td>0.6, 0.4</td>
<td>0.753</td>
</tr>
<tr>
<td>3</td>
<td>0.34, 0.33, 0.33</td>
<td>0.268</td>
</tr>
<tr>
<td></td>
<td>0.4, 0.35, 0.25</td>
<td>0.410</td>
</tr>
<tr>
<td></td>
<td>0.5, 0.3, 0.2</td>
<td>0.664</td>
</tr>
<tr>
<td>4</td>
<td>0.26, 0.25, 0.25, 0.24</td>
<td>0.214</td>
</tr>
<tr>
<td></td>
<td>0.4, 0.3, 0.2, 0.1</td>
<td>0.512</td>
</tr>
<tr>
<td></td>
<td>0.5, 0.3, 0.1, 0.1</td>
<td>0.708</td>
</tr>
<tr>
<td>5</td>
<td>0.21, 0.2, 0.2, 0.2, 0.19</td>
<td>0.157</td>
</tr>
<tr>
<td></td>
<td>0.3, 0.2, 0.2, 0.2, 0.1</td>
<td>0.360</td>
</tr>
<tr>
<td></td>
<td>0.35, 0.2, 0.15, 0.15, 0.15</td>
<td>0.506</td>
</tr>
</tbody>
</table>
A second thing to note from Table 1 is how plurality rule performs where voters are just slightly more likely to choose the correct option than incorrect ones. Where each voter’s probability of choosing the correct option from among \( k \) options is just over \( 1/k \) and the probability of choosing incorrect ones just under that, the probability of the correct option being the plurality winner increases much more slowly with increases in the size of the electorate. Compare this with the standard two-option Condorcet jury result in the case in which each voter is just over \( 1/2 \) likely to choose the correct option: as Table 1 shows, it takes over a thousand voters before the probability of the correct option being the plurality (majority) winner is 0.737, where each voter has only a probability of 0.51 of voting for the correct outcome. Similarly for the many-option case: the probability of the correct option being the plurality winner increases much more slowly where the probability of each voter being correct is near \( 1/k \), compared to cases where the probability of each voter choosing the correct outcome is even just a little higher. But even in these "worst-case" scenarios, the movement of the probability figures is clearly in the desired direction; and, as the size of the electorate increases, the probability of the correct option being the plurality winner will eventually approach certainty.

IV. Comparing Truth-Trackers

The plurality rule is not the unique truth-tracker in the \( k \)-option case. Condorcet himself pointed to others, in passages in his *Essai* immediately following his development of the jury theorem itself.\(^{39}\) Contemporary social

choice theorists tend to see a sharp disjunction at this point in his text\(^{40}\); and in
terms of technique and methodology there certainly is.\(^{41}\) But Condorcet's
own concerns remained resolutely epistemic throughout; and those grounds
(rather than the proceduralist ones more standardly attributed to him by
many contemporary social-choice interpreters) are the ones on which
Condorcet himself proceeds to recommend what has become famous as the
"Condorcet winner" criterion based on pairwise comparisons for \(k\)-option
cases.\(^{42}\)

Condorcet's own analysis at this point the *Essai* is notoriously
opaque.\(^{43}\) In consequence, it lay largely neglected for most of the intervening
two centuries. Duncan Black saw what Condorcet was up to, but he was
unable to elucidate it in a way that seized the broader attention of democratic
theorists in the same way that his exposition of the jury theorem did.\(^{44}\) More
recently, Young's results are effectively a restatement of Condorcet's analysis
in modern statistical terms; but that restatement, too, seems to have largely

---

\(^{40}\) As Arrow (1963, pp. 94-5) says, "Condorcet has really two different approaches. In the
one most in line with subsequent developments, as well as with Borda's work, the chief
contribution has been what might be termed the Condorcet criterion, that a candidate
who receives a majority as against each other candidate should be elected. This
implicitly accepts the view of what I have termed the independence of irrelevant
alternatives. ... The second approach is closely related to the theory of juries which
Condorcet and others were studying. Here the implication is rather hat voters are
judges of some truth rather than expressing their own preferences." McLean and Hewitt
(1994, 34-5) echo, "The theory of voting in the [Condorcet] *Essai* of 1785 is really two
theories; furthermore, they are inconsistent. When forced to make a choice between the
probabilistic theory and the social choice theory, Condorcet opts for the latter, even
though the former occupies the vast majority of the book."

\(^{41}\) "For whatever reason, Cordorcet abruptly changed course at this point in the *Essai*.... He
abandoned the statistical framework that he had so painstakenly built up and fell
back on a more 'straightforward' line of reasoning [by pairwise comparisons and such
like]. In so doing he opened up a whole new approach that has had enormous influence
Black 1958, pp. 168-71; McLean and Hewitt 1994, pp. 8383-40. The shift in gears is
signalled clearly by Condorcet (1785/1972, pp. lxiii-lxv, 122-4) himself.


\(^{43}\) One commentator says, "It is quite hopeless to find out what Condorcet meant," at this
point; another that, "The obscurity and self-contradiction are without any parallel, so
far as our experience of mathematical works extends..." (both quoted in Young 1988, p.
1234).

\(^{44}\) Black 1958, pp. 169-7. Cf. the take-up of Black's discussion of the jury theorem in Barry
(1964 , pp. 9-14; 1965, Appendix A, pp. 292-3) and, through him, all the others cited at
footnote 25 above.
escaped the notice of more philosophical commentators on epistemic democracy.45

To get a grip on Condorcet’s approach, go back to the two-option case. There, Condorcet knew from his jury theorem that majority rule was the best truth-tracker. But then the problem was what the most ”natural” way to extend majority rule beyond the two-option case. As the ”scare quotes” suggest, there is no uniquely ”natural” extension. Majority rule is a special case, for \( k=2 \), of an (indefinitely) large number of plausible decision rules that might be applied in the case of \( k>2 \). Fixing majority rule as the appropriate decision procedure for the two-option case still logically underdetermines our choice of an appropriate decision procedure for the many-option case.46 One alternative is the plurality rule, as just discussed. But Condorcet (and Borda before him) had already exposed the apparent irrationalities of that rule where \( k>2 \), so he did not consider it a viable option worthy of further discussion at this point. Instead, he examined another of the ”natural” extensions of majority rule to the many-option case: pairwise comparison. The attraction of that rule, perhaps, was that Condorcet inferred from his jury theorem that in binary choice situations (which each of the pairwise choices are, of course) whichever option is chosen by a majority is most likely to be right, assuming choosers are more likely to be right than wrong on average. His thought seems to have been that, if each of the pairwise choices is likely to be correct, then the overall outcome of a series of such choices is likely to be correct too.

To Condorcet’s frustration, the logic did not quite privilege his pairwise comparisons uniquely. Everything turns out to depend upon how much more than half-likely voters were to be correct. If they were much

---

45 Young 1988, pp. 1238 ff.; 1995. This paper is well-known among Condorcet scholars, of course (McLean and Hewitt 1994, pp. 38-40), and technical social choice theorists. Among democratic theorists more generally, however, we have noticed only one passing reference to Young’s paper (Miller 1992, p. 66 n. 28); and that failed to emphasize what we here regard as really distinctive about that paper.

46 Quine 1961.
more likely to be correct in each pairwise choice (i.e. if their competence is close to 1), then the option most likely to be the correct one is the winner under Condorcet’s pairwise method. But if voters were only barely more than half-likely to be correct in each pairwise choice (i.e. if their competence is close to 0.5), then the option most likely to be the correct one is the winner under the Borda count. This, in a nutshell, is the result that Condorcet discovered and that Young has proven more formally.47

The history, however interesting, is neither here nor there. Our purpose in recounting the tale is merely to remind ourselves that several decision rules might have considerable epistemic merit in the \(k\)-option case. One, as we have shown above, is the plurality rule. Others, as shown by Young and Condorcet himself, include pairwise comparisons and the Borda count — two of the decision rules most cherished among contemporary procedural democrats. The informational requirements of the latter sorts of rules are, of course, much greater: they need complete rankings of all options from all voters, whereas plurality rules need only know each voter’s first choice among all the options. But given that extra information, those other rules track the truth too — in fact, better than plurality rule.

The actual numbers matter, though. In the discussion of Section II, it would have been cold comfort to epistemic democrats that the plurality rule is a good truth-tracker, just so long as the electorate is sufficiently large — if "sufficiently large" had turned out to be some preposterously large number (billions of billions, say). In the present discussion, it would be similarly cold comfort to epistemic democrats that some particular decision rules track the

47 Condorcet himself did not pursue the matter further, shifting analytic gears at this point in the Essai, abandoning his probability calculus of truth and error, and turning instead to the sorts of analyses for which he is famous among today’s social choice theorists. One reason might have been sheer personal spite: Borda was Condorcet’s great nemesis in the Academy of Science, and some commentators suggest that he dropped the whole line of enquiry rather than supporting the proposals of his great adversary (Young 1988, p. 1238; McLean and Hewitt 1994, p. 40).
truth better than others, if even the best truth-tracker turns out to track the truth abysmally badly, by any objective standards.

In the computational exercise that follows, we set out one plausible procedure for calculating the probabilities that each of the standardly-discussed decision rules will pick the epistemically correct outcome, under varying assumptions about the probabilities that each voter has of choosing the correct (and each incorrect) option and about the number of voters. These calculations of course represent only a small sample of all possible such combinations; as such, they strictly speaking "prove" nothing. Still, they are illustrative, and the general outlines of the picture they sketch soon enough become tolerably clear.

To generate these probability calculations for each decision rule in the $k$-option case, we require some way of moving from (a) assumptions about the probability that each voter will choose each option (as set out in the framework of our plurality jury theorem) to (b) inferences about the frequencies with which voters can be expected to harbor particular "complete orderings" of preferences over all options. To move from (a) (the narrower informational framework, in which the plurality jury theorem holds) to (b) (the richer informational framework, in which rules like Condorcet's or Borda's are applicable), we employ a specific heuristic that seems to us particularly appealing. But it is of course only one among many possible such heuristics for moving from (a) to (b). So in that sense, too, our calculations here are merely illustrative, no more.

Details of our heuristic are set out in Appendix 2. Here, suffice it to say that we start with a set of probabilities, as in the second column of Table 1, representing the probabilities each individual has of choosing each option. We let those probabilities dictate the probability with which each of those respective options will appear as the first-choice option in each person's preference ordering; for each possible first-choice option, we then let the
relative probabilities associated with each of the remaining options dictate the probability of each of these options' appearing as the second-choice option in the same preference ordering; and so on until all places in the preference ordering have been allocated. The probability with which any given preference ordering will be expected is adduced in this way from the product of the probabilities of filling each of the places with the relevant options in this fashion.

As we say, this is not the only way of proceeding from individuals' probability profiles to probabilities of overall preference rankings. But it has a certain surface plausibility about it. True, our procedure does not allow for the possibility of incomplete, intransitive or inconsistent preference orderings at the individual level. But in this respect, our procedure maps a central feature of how electoral systems themselves actually work, when evoking full preference orderings from people. There, just as in our procedure, voters are typically required to rank options by assigning exactly one rank to each option.

Using this heuristic, we generate (by stipulation) probabilities of each voter holding each of various preference orderings from information about probabilities of each voter supporting each of various options. Given that information, we can then calculate the probabilities with which each option would win under each of various social decision rules — not just the plurality rule, but also the Borda count, the Condorcet pairwise comparison criterion, and the Hare and Coombs systems. The probability that particularly interests us, in the present context, is of course the probability that the outcome we have stipulated as "epistemically correct" will emerge under each of those decision procedures.

Table 2 reports the probability that correct outcome emerge from various social decision rules, under various scenarios (different numbers of options, different probabilities of each voter supporting each) and for
electorates of varying size. In this, we compare the performance of five social
decision rules: the plurality rule; the pairwise-Condorcet rule; the Borda
count; the Hare system; and the Coombs system. To keep the computations
manageable, we have restricted our attention to cases where $k \leq 3$ and to cases
where the size of the electorate is 71 or smaller. To simplify the presentation
of our results, we report the probabilities of the correct option emerging as
the outcome under those various decision rules for just a few examples
involving electorates of different sizes (11, 31, 51 and 71) and a few selected
probabilities of each voter choosing correct and incorrect outcomes.

[Table 2 about here]
<table>
<thead>
<tr>
<th></th>
<th>Scenario 1:</th>
<th>Scenario 2:</th>
<th>Scenario 3:</th>
<th>Scenario 4:</th>
<th>Scenario 5:</th>
<th>Scenario 6:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$k = 2$</td>
<td>$k = 3$</td>
<td>$k = 3$</td>
<td>$k = 3$</td>
<td>$k = 3$</td>
<td>$k = 3$</td>
</tr>
<tr>
<td></td>
<td>$p_1 = 0.51$</td>
<td>$p_1 = 0.6$</td>
<td>$p_1 = 0.51$</td>
<td>$p_1 = 0.40$</td>
<td>$p_1 = 0.34$</td>
<td>$p_1 = 0.335$</td>
</tr>
<tr>
<td></td>
<td>$p_2 = 0.49$</td>
<td>$p_2 = 0.3$</td>
<td>$p_2 = 0.25$</td>
<td>$p_2 = 0.30$</td>
<td>$p_2 = 0.33$</td>
<td>$p_2 = 0.3325$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$p_3 = 0.1$</td>
<td>$p_3 = 0.24$</td>
<td>$p_3 = 0.30$</td>
<td>$p_3 = 0.33$</td>
<td>$p_3 = 0.3325$</td>
</tr>
<tr>
<td><strong>plurality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n=11$</td>
<td>0.527</td>
<td>0.821</td>
<td>0.699</td>
<td>0.419</td>
<td>0.268</td>
<td>0.256</td>
</tr>
<tr>
<td>$n=31$</td>
<td>0.545</td>
<td>0.957</td>
<td>0.915</td>
<td>0.581</td>
<td>0.322</td>
<td>0.302</td>
</tr>
<tr>
<td>$n=51$</td>
<td>0.557</td>
<td>0.988</td>
<td>0.972</td>
<td>0.666</td>
<td>0.338</td>
<td>0.311</td>
</tr>
<tr>
<td>$n=71$</td>
<td>0.567</td>
<td>0.996</td>
<td>0.991</td>
<td>0.726</td>
<td>0.344</td>
<td>0.313</td>
</tr>
<tr>
<td><strong>pairwise Condorcet</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n=11$</td>
<td>0.527</td>
<td>0.877</td>
<td>0.805</td>
<td>0.514</td>
<td>0.327</td>
<td>0.312</td>
</tr>
<tr>
<td>$n=31$</td>
<td>0.545</td>
<td>0.973</td>
<td>0.959</td>
<td>0.652</td>
<td>0.339</td>
<td>0.313</td>
</tr>
<tr>
<td>$n=51$</td>
<td>0.557</td>
<td>0.993</td>
<td>0.991</td>
<td>0.740</td>
<td>0.348</td>
<td>0.315</td>
</tr>
<tr>
<td>$n=71$</td>
<td>0.567</td>
<td>0.998</td>
<td>0.998</td>
<td>0.802</td>
<td>0.356</td>
<td>0.317</td>
</tr>
<tr>
<td><strong>Borda</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n=11$</td>
<td>0.527</td>
<td>0.864</td>
<td>0.805</td>
<td>0.505</td>
<td>0.316</td>
<td>0.301</td>
</tr>
<tr>
<td>$n=31$</td>
<td>0.545</td>
<td>0.976</td>
<td>0.967</td>
<td>0.666</td>
<td>0.344</td>
<td>0.318</td>
</tr>
<tr>
<td>$n=51$</td>
<td>0.557</td>
<td>0.995</td>
<td>0.994</td>
<td>0.760</td>
<td>0.360</td>
<td>0.326</td>
</tr>
<tr>
<td>$n=71$</td>
<td>0.567</td>
<td>0.999</td>
<td>0.999</td>
<td>0.823</td>
<td>0.371</td>
<td>0.330</td>
</tr>
<tr>
<td><strong>Hare</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n=11$</td>
<td>0.527</td>
<td>0.875</td>
<td>0.801</td>
<td>0.527</td>
<td>0.352</td>
<td>0.338</td>
</tr>
<tr>
<td>$n=31$</td>
<td>0.545</td>
<td>0.973</td>
<td>0.955</td>
<td>0.655</td>
<td>0.365</td>
<td>0.341</td>
</tr>
<tr>
<td>$n=51$</td>
<td>0.557</td>
<td>0.993</td>
<td>0.989</td>
<td>0.760</td>
<td>0.369</td>
<td>0.338</td>
</tr>
<tr>
<td>$n=71$</td>
<td>0.567</td>
<td>0.998</td>
<td>0.998</td>
<td>0.832</td>
<td>0.382</td>
<td></td>
</tr>
<tr>
<td><strong>Coombs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n=11$</td>
<td>0.527</td>
<td>0.877</td>
<td>0.816</td>
<td>0.539</td>
<td>0.354</td>
<td>0.338</td>
</tr>
<tr>
<td>$n=31$</td>
<td>0.545</td>
<td>0.973</td>
<td>0.962</td>
<td>0.675</td>
<td>0.368</td>
<td>0.342</td>
</tr>
<tr>
<td>$n=51$</td>
<td>0.557</td>
<td>0.993</td>
<td>0.991</td>
<td>0.755</td>
<td>0.372</td>
<td>0.339</td>
</tr>
<tr>
<td>$n=71$</td>
<td>0.567</td>
<td>0.998</td>
<td>0.998</td>
<td>0.835</td>
<td>0.385</td>
<td>0.346</td>
</tr>
</tbody>
</table>
Definitions:\n
*Plurality rule:* "Choose the candidate who is ranked first by the largest number of voters."

*Condorcet pairwise criterion:* "Choose the candidate [if unique] who defeats [or at least ties with] all others in pairwise elections using majority rule."

*Borda count:* "Give each of the $m$ candidates a score of 1 to $k$ based on the candidate's ranking in a voter's preference ordering; that is, the candidate ranked first receives $m$ points, the second one $k-1$, ..., the lowest-ranked candidate one point. The candidate [if unique] with the highest number of points is declared the winner."

*Hare system:* "Each voter indicates the candidate he ranks highest of the $k$ candidates. Remove from the list of candidates the one [or in case of ties, ones] ranked highest by the fewest voters. Repeat the procedure for the remaining $k-1$ candidates. Continue until only [at most] one candidate remains. Declare this candidate [if any] the winner."

*Coombs system:* "Each voter indicates the candidate he ranks lowest of the $k$ candidates. Remove from the list of candidates the one [or in case of ties, ones] ranked lowest by the most voters. Repeat the procedure for the remaining $k-1$ candidates. Continue until only [at most] one candidate remains. Declare this candidate [if any] the winner."

\textsuperscript{48} All definitions are from Mueller 1989, pp. 112-3.
Before proceeding to any more detailed commentary on Table 2, one important thing to say about all the calculations within it is this. The probabilities reported in the cells of that table represent the probability with which the correct outcome will be uniquely chosen by each decision rule. Decision rules can fail to do so in either of two ways. One is by choosing the wrong outcome. Another is by choosing no outcome, or anyway none uniquely. Sometimes, for example, there simply is no Condorcet winner; where there is not, we count that as a failure. And sometimes decision rules produce no unique winner; we count indecisiveness, in cases of "ties," as a failure as well. The probability statistics in Table 2 thus reflect decisiveness as well as correctness per se.49

As we have said, all these decision rules are extensionally equivalent to one another in two-option case. That is shown in Table 2, Scenario 1. That scenario represents the "standard" Condorcet jury theorem finding, in its classical \( k=2 \) form. And it serves as a benchmark against which the epistemic performance of other decision rules in \( k>2 \) cases can be compared.

Where the probability of each voter choosing the correct option remains at 0.51, but the number of options increases from \( k=2 \) to \( k=3 \), the probability of the correct outcome being chosen is greatly increased over that of the correct outcome being chosen in the two-option case. We have noted that already in connection with plurality voting, in our discussion of Table 1. What we see from Table 2, when comparing Scenarios 1 and 3, is that that is true (indeed, even more true) of all of the other standard social decision rules as well.

---

49 Casual inspection of sample calculations may seem to suggest that that might be why the Borda rule overtakes the Condorcet pairwise rule, where \( n \) is large. It is clearly the reason why the Hare and Coombs rules seem to become "less reliable" for \( n=51 \) than for \( n=31 \). Those rules can be indecisive (i.e., result in ties) in the \( k=3 \) case only when \( n \) is a multiple of 2 or 3; and that affects the case of \( n=51 \) uniquely among the values of \( n \) for which Table 2 calculates these probabilities.
Where the probability of each voter choosing the correction option drops to just over $1/k$, and the probabilities of choosing each of the wrong ones to just under that, all of these decision rules will obviously require large electorates (larger than the computing power available to us allows us to analyze) in order to achieve any very high degree of epistemic strength. That is seen clearly from Scenarios 5 and especially 6 in Table 2. But what is clear from the computations we have been able to perform is that the epistemic strength of each of the decision rules increases with the size of the electorate. If Young's result (based though it is on rather different assumptions) can be applied, then there is every reason to believe that Condorcet or Borda rules, anyway, will if anything exceed the high epistemic strength found in Table 1 for the plurality rule.

There are many odd and interesting small differences among decision rules revealed in Table 2. Some of them might be quirks or artefacts of our particular methodology for calculating the probabilities. Others might reflect deeper facts about the decision rules in question.

The principal things we want to point out about Table 2, however, are not the differences but rather the broad similarities among all these decision rules. Particularly where the size of the electorate is at all large (51 or over, say), each of these decision rules is pretty nearly as good a truth-tracker as any other. Even in the worst case, in Scenario 6, the epistemic strength of the worst decision rule (plurality) is only a few percentage points worse than the best (Hare or Coombs), in electorates of 51 or more.

That is the first "big" conclusion we would draw from Table 2. Any of these standard decision rules is pretty much as good as any other, on epistemic grounds. We are at liberty to choose among them, according to

---

50 The exception in Table 2 is with the Hare and Coombs rules, in Scenario 6. The explanation for that exception is provided in the preceding footnote.

51 Such as the apparent "reversal" with Hare and Coombs rules, discussed in the preceding note.
their varying proceduralist merits, pretty much without fear of any epistemic consequences.

The second "big" conclusion we would draw from Table 2 is this. All of these standard decision rules have great epistemic merits, at least whenever the electorate is reasonably large. These merits are greatest where the probability of each voter choosing the correct outcome is substantially larger than $1/k$ (Scenarios 2 and 3). But these merits are still great, at least where the electorate is at all large (over 51, say), even where the probability of each voter choosing the correct outcome is much nearer the probabilities of each voter choosing incorrect ones (Scenario 4). It is only where the probability of each voter choosing correctly is just barely over $1/k$, and of choosing each incorrect option is just under that, that very large electorates will be required to yield really reliable results (Scenarios 5 and 6). But even there, with a realistically large electorate (the size of a city, say), epistemic strength will grow high. And all of that seems broadly speaking true of all the standard decision rules ordinarily canvassed.

IV. Conclusions

Social choice theorists and electoral reformers squabble endlessly over what is the "best" democratic decision rule from a procedural point of view. Here we have shown that we can afford to be relatively relaxed about that choice from an epistemic point of view. Some social decision rules (Coombs and Hare) seem to be marginally better truth-trackers than others. But when the electorate is even remotely large, all of the standardly-discussed decision rules (including even the plurality rule) are almost equally good truth-trackers. There is little to choose among them, on epistemic grounds.
Furthermore, all of them are good truth-trackers — insofar, of course, as there are any "truths" for politics to track at all.\textsuperscript{52} Just how good they all are depends on the size of the electorate and the reliability of electors. But even in the worst-case scenarios, it takes only city-sized electorates to allow us to be highly confident that the epistemically correct outcome emerges under any of the standard democratic decision rules (just so long as we can be minimally confident in the reliability of individual voters). In short, democracy in any of its standard forms is potentially a good truth-tracker: it can always hope to claim that epistemic merit, whatever other procedural merits any particular version of it might also manifest.

Thus, we have not so much settled these standing controversies in democratic theory as circumvented them.\textsuperscript{53} Proceduralists of the social-choice sort who are enamored of the axiomatic merits of the Condorcet pairwise rule, for example, may feel free to recommend that rule on democratic-proceduralist grounds, without fear of any great epistemic costs. Old-fashioned democratic proceduralists who are anxious that people be governed by rules that they can understand, and who are thus attracted to the plurality rule by reasons of its sheer simplicity and minimal informational requirements, may feel almost equally free to recommend that rule without any great epistemic costs. Democracy has potentially great epistemic merits, in any of its many forms.

\textsuperscript{52} Cf. Black 1958, p. 163; Miller 1992, p. 56.
\textsuperscript{53} Assuming, of course, our results stand up to further, more detailed scrutiny of the sort we indicated above is needed (at footnotes 36-7). Specifically, strategic manipulation might be much more of a risk in many-option cases, and different decision rules might be differentially vulnerable to it. Further investigation of those issues is clearly required. Pending those further investigations, however, these represent our present, tentative conclusions.
Appendix 1: A Simple k-option Condorcet Jury Model

Suppose that there are $n$ voters/jurors, that there are $k$ options, $x_1, x_2, ..., x_k$ (where $k>1$), and that each voter/juror has independent probabilities $p_1, p_2, ..., p_k$ of voting for $x_1, x_2, ..., x_k$ as his/her first choice, respectively (where $\sum p_i = 1$).

Let $X_1, X_2, ..., X_k$ be the random variables whose values are the numbers of first-choice votes (out of a total of $n$ votes) cast for $x_1, x_2, ..., x_k$ respectively.

The joint distribution of $X_1, X_2, ..., X_k$ is a multinomial distribution with the following probability function:

$$P(X_1=n_1, X_2=n_2, ..., X_k=n_k) = \frac{n!}{n_1! n_2! ... n_k!} p_1^{n_1} p_2^{n_2} ... p_k^{n_k}.$$  

For each $i$, the mean of $X_i$ is $\mu_i = np_i$, the variance of $X_i$ is $\sigma_i^2 = np_i(1-p_i)$, and, for each $i$ and $j$ ($i \neq j$), the covariance of $X_i$ and $X_j$ is $\sigma_{ij}^2 = -np_ip_j$.

**Proposition 1.** For each $i$, the probability that $x_i$ will win under plurality voting is

$$P(\text{for all } j \neq i, X_i > X_j) = \frac{n!}{\sum_{n_1, n_2, ..., n_k: \sum p_i=n & \text{for all } j \neq i, n_j \neq n_i} p_1^{n_1} p_2^{n_2} ... p_k^{n_k}} n_1! n_2! ... n_k!.$$  

In particular, note that since the joint distribution of $X_1, X_2, ..., X_k$ is multinomial, if $p_i$ is greatest among $p_1, p_2, ..., p_k$, then so is $P(\text{for all } j \neq i, X_i > X_j)$ (i.e. the probability that $x_i$ will win under plurality voting) among the probabilities corresponding to $x_1, x_2, ..., x_k$.

**Proposition 2.** Suppose, for a fixed $i$, we have, for all $j \neq i$, $p_i > p_j$. Then the probability that $x_i$ will win under plurality voting tends to 1 as $n$ tends to infinity, i.e. $P(\text{for all } j \neq i, X_i > X_j) \to 1$ as $n \to \infty$. 


Sketch proof. Consider the vector of random variables $X^* = \langle X^*_1, X^*_2, \ldots, X^*_k \rangle$, where, for each $i$, $X^*_i := X_i/n$. The joint distribution of the $X^*_i$ is a multinomial distribution with mean vector $\mu = \langle p_1, p_2, \ldots, p_k \rangle$ and with variance-covariance matrix $\Sigma = (s_{ij})$, where, for each $i$ and each $j$, $s_{ij} = p_i(1-p_i)$ if $i=j$ and $s_{ij} = -p_ip_j$ if $i \neq j$. By the central limit theorem, for large $n$, $\sqrt{n}(X^*-\mu)$ has an approximate multivariate normal distribution $N(0, \Sigma)$. This implies that, for large $n$, $X^*-\mu$ has an approximate multivariate normal distribution $N(0, 1/n \Sigma)$.

Let $f_n : \mathbb{R}^k \to \mathbb{R}$ be the corresponding density function for $X^*-\mu$. Using this density function, the probability that option $x_i$ will win under plurality voting is given by $P(\text{for all } j \neq i, X^*_i > X^*_j) = \int_{W_i} f_n(t) dt$, where $W_i := \{ t = \langle t_1, t_2, \ldots, t_k \rangle \in \mathbb{R}^k : \text{for all } j \neq i, p_i + t_i > p_j + t_j \}$. Since, by assumption, for all $j \neq i$, $p_i > p_j$, there exists an $\varepsilon > 0$ such that $S_{\varepsilon} \subseteq W_i$, where $S_{\varepsilon}$ is a sphere around $0$ with radius $\varepsilon$. Then, since $f_n$ is nonnegative, $\int_{S_{\varepsilon}} f_n(t) dt \geq \int_{S_{\varepsilon}} f_n(t) dt$. But, as $f_n$ is the density function corresponding to $N(0, 1/n \Sigma)$, $\int_{S_{\varepsilon}} f_n(t) dt \to 1$ as $n \to \infty$, and the desired result follows. Q.E.D.
Appendix 2: A Simple Heuristic for Deriving Probabilities over Preference Orderings from Probabilities over Single Votes

The model introduced in appendix 1 is suitable for assessing the epistemic qualities of plurality voting and, more generally, of voting procedures whose input is a single vote, or most preferred option, for each voter/juror. To assess the epistemic qualities of voting procedures whose input is a complete preference ordering (rather than just a single vote or most preferred option) for each voter/juror, more information is required. We will extend our model as follows. Given the \( k! \) logically possible strict preference orderings, \( P_1, P_2, \ldots, P_{k!} \) over the \( k \) options, \( x_1, x_2, \ldots, x_k \) we will assume that each voter/juror has independent probabilities \( p_1^*, p_2^*, \ldots, p_{k!}^* \) of submitting \( P_1, P_2, \ldots, P_{k!} \) as his/her preference ordering, respectively (where \( \sum p_i = 1 \)).

Now let \( X_1^*, X_2^*, \ldots, X_{k!}^* \) be the random variables whose values are the numbers of voters/jurors submitting the orderings \( P_1, P_2, \ldots, P_{k!} \) respectively.

Again, the joint distribution of \( X_1^*, X_2^*, \ldots, X_{k!}^* \) is a multinomial distribution with the following probability function:

\[
P(X_1^* = n_1, X_2^* = n_2, \ldots, X_{k!}^* = n_{k!}) = \frac{n!}{n_1! n_2! \ldots n_{k!}!} p_1^{n_1} p_2^{n_2} \ldots p_{k!}^{n_{k!}}.
\]

Given any criterion for determining a winning option (such as the pairwise Condorcet, Borda, Hare, Coombs and of course plurality criteria), we can then use this probability function to compute, for each \( i \), the probability that option \( x_i \) will win under the given criterion.

To compare the epistemic qualities of plurality voting with those of voting procedures whose input is a complete preference ordering for each voter/juror, we use a simple heuristic for deriving the probabilities \( p_1^*, p_2^*, \ldots, p_{k!}^* \) associated with the preference orderings \( P_1, P_2, \ldots, P_{k!} \) from the given probabilities \( p_1, p_2, \ldots, p_k \) associated with the options \( x_1, x_2, \ldots, x_k \).
In the original $k$-option jury model, an individual voter/juror's vote is effectively modeled as a single draw from an urn with a proportion of $p_1, p_2, \ldots, p_k$ balls of types $x_1, x_2, \ldots, x_k$, respectively. Similarly, in the new model, an individual voter/juror's strict preference ordering over $k$ options will be modeled as a sequence of $k$ draws (corresponding to the $k$ ranks in the preference ordering) from an urn with an initial proportion of $p_1, p_2, \ldots, p_k$ balls of types $x_1, x_2, \ldots, x_k$, respectively, and where after each draw all balls of the type drawn are removed, so that eventually, in the $k$-th (and last) draw only one type of balls is left in the urn. Now the probability associated with an ordering $P_i$ is simply the probability that, in this urn model, the options are drawn in precisely the order in which they are ranked by the ordering $P_i$. Formally, if $P_i$ is the ordering $x_{i_1} > x_{i_2} > \ldots > x_{i_k}$, the probability associated with $P_i$ is simply

\[
p_{i_1} \cdot p_{i_2} \cdot p_{i_3} \cdot \ldots \cdot p_{i_{k-1}} \cdot 1.
\]

To illustrate, table 3 lists the probabilities associated with all logically possible orderings derived from the probabilities associated with single votes in the three-option scenarios of table 2.
Table 3: Probabilities over preference orderings derived from probabilities over single votes

<table>
<thead>
<tr>
<th>p option x₁ (correct)</th>
<th>Scenario 2: k=3</th>
<th>Scenario 3: k=3</th>
<th>Scenario 4: k=3</th>
<th>Scenario 5: k=3</th>
<th>Scenario 6: k=3</th>
</tr>
</thead>
<tbody>
<tr>
<td>p option x₂</td>
<td>0.6</td>
<td>0.51</td>
<td>0.40</td>
<td>0.34</td>
<td>0.335</td>
</tr>
<tr>
<td>p option x₃</td>
<td>0.3</td>
<td>0.25</td>
<td>0.30</td>
<td>0.33</td>
<td>0.3325</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>0.24</td>
<td>0.30</td>
<td>0.33</td>
<td>0.3325</td>
</tr>
<tr>
<td>p x₁&gt;x₂&gt;x₃</td>
<td>0.450</td>
<td>0.260</td>
<td>0.200</td>
<td>0.170</td>
<td>0.1675</td>
</tr>
<tr>
<td>p x₁&gt;x₃&gt;x₂</td>
<td>0.150</td>
<td>0.250</td>
<td>0.200</td>
<td>0.170</td>
<td>0.1675</td>
</tr>
<tr>
<td>p x₂&gt;x₁&gt;x₃</td>
<td>0.257</td>
<td>0.170</td>
<td>0.171</td>
<td>0.168</td>
<td>0.1675</td>
</tr>
<tr>
<td>p x₂&gt;x₃&gt;x₁</td>
<td>0.043</td>
<td>0.080</td>
<td>0.129</td>
<td>0.168</td>
<td>0.1675</td>
</tr>
<tr>
<td>p x₃&gt;x₁&gt;x₂</td>
<td>0.067</td>
<td>0.161</td>
<td>0.171</td>
<td>0.162</td>
<td>0.1675</td>
</tr>
<tr>
<td>p x₃&gt;x₂&gt;x₁</td>
<td>0.033</td>
<td>0.079</td>
<td>0.129</td>
<td>0.162</td>
<td>0.1675</td>
</tr>
</tbody>
</table>
References


Bartholdi, J.; C. A. Tovey; and M. A. Trick. 1989. Voting schemes for which it can be difficult to tell who won the election. Social Choice and Welfare, 6 (#2): 157-65.


Grofman, Bernard; Guillermo Owen; and Scott L. Feld. 1983. Thirteen theorems in search of the truth. Theory & Decision, 15: 261-78.


