version 9 January 2001 – to appear in Journal of Political Philosophy

Epistemic Democracy: Generalizing the Condorcet Jury Theorem*

CHRISTIAN LIST

Nuffield College, Oxford

and

ROBERT E. GOODIN

Philosophy and Social & Political Theory, Australian National University

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."[1] 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.[2]

For procedural democrats, the aim of democracy is instead to embody certain procedural virtues.[3] 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, democracy is not about tracking any "independent truth of the matter"; instead, the goodness or rightness of an outcome is wholly constituted by the fact of its having emerged in some procedurally correct manner.[4]

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 correct outcome is most likely to win a majority of votes.[5] 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.[6] Although the many different rules that different procedural democrats recommend (Condorcet pairwise comparisons, the Borda count, the Hare or Coombs systems, etc.[7]) 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, converge on the majority winner.[8]

This happy coincidence is confined to the two-option case, however. Where there are three or more options on the table, recommendations of the different strands of democratic theory diverge.[9] 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.[10] 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[11]) that the 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 that it is uniquely preferred, epistemically.[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 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.

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. There are intermediate and mixed versions as well, but here we shall be concerned with each type of theory in its most extreme, pure form.

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. A pure 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.[14] 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.

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.

Where that procedure cannot itself be perfectly implemented as the social decision rule, advocates of the procedural approach must choose among the array of procedures that are actually available whatever one of them best approximates the right-making dictates of that perfect procedure.[17]

What social procedures, if any, should be regarded as constitutive in this way of best or right outcomes is of course a highly contentious issue – and one which we do not here propose to resolve.[18]

Let us just offer an illustrative list of a few of the very different sorts of answers to this question that have been proposed from time to time:

- Democratic proceduralism most narrowly construed highlights the properties of aggregation procedures, that is, those procedures by which votes or individual "preference" input are transformed into social decisions. Proceduralists of that bent have argued, with increasing precision and formality over the past couple of centuries, that we should employ aggregation procedures which make social decisions systematically responsive to the preferences expressed by individual voters or decision-makers; and they have increasingly come to insist that that should be understood as meaning that the procedures 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[19]). At the most formal end of the spectrum, democratic proceduralists have recently added various axiomatic desiderata to the list of procedural criteria. They typically specify a set of (normative) minimal conditions that any acceptable aggregation procedure should satisfy[20]; and they then determine what aggregation procedures, if any, satisfy these conditions.[21]
- Democratic proceduralism more broadly construed highlights a whole set of institutional and political arrangements relevant to social decisions, particularly the question of what political processes should lead to social decisions, what role political communication should play,

> who should participate in social decision processes, and in what form and how regularly such participation should take place. Democratic proceduralists of this broader variety have 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 publicly transparent (which is perhaps the main thing the "first past the post" plurality rule has going for it, procedurally[22]); that social decisions should be preceded by certain processes of reasoned political deliberation and communication, and that people affected by a decision ought be heard; and also that social decision procedures should be practically viable and implementable at acceptable costs.[23]

Here we shall be concerned with proceduralism most narrowly construed. That is to say, we shall (in Section III below) simply be examining different procedures for aggregating votes into an overall social decision.

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

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.[24]

Extrapolating from juries to electorates more generally, that result constitutes the jewel in the crown of epistemic democrats, many of whom offer it as powerful evidence of the truth-tracking merits of majority rule.[25] 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.[26] 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.[27] The effects of strategic voting in a Condorcet jury context have also been studied extensively, showing mixed results.[28] 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 *pairwise* votes between them.[29]) The choice is thus typically 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.[30]

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).[31] 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.[32]

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 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.[33]

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.[34] Specifically, we show the following:

Proposition 1. Suppose there are k options and that each voter/juror has independent probabilities $p_1, p_2, ..., p_k$ of voting for options 1, 2, ..., 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.

Proposition 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 proof is contained in Appendix 1.[35] But informally, what drives the result is just the law of large numbers. Think about tossing a fair coin, which has a $p_{heads}=p_{tails}=0.50$ chance of landing either heads or tails. In a small number of tosses (say 100), the actual numbers of heads and tails might well be 60:40, which is a considerable deviation from the expected proportions of 50

percent and 50 percent. But in a larger number of tosses (say 1000), the actual numbers might be 530:470, which is closer to the expected proportions of 50 percent and 50 percent. The point of the law of large numbers is that, although *absolute* deviations from the expected numbers still increase as the number of trials increases, those absolute deviations are a decreasing *proportion* of the total as the number of trials increases.

So too with voters. Among voters who are each p=0.51 likely to vote for a proposition, the statistically expected distribution of votes would be 51 percent in favor and 49 percent opposed to the proposition. Similarly, among voters who are p_1=0.40 likely to vote for option 1 and p_2=p_3=0.30 likely to vote for options 2 and 3 respectively, the statistically expected distribution of votes would be 40 percent for option 1 and 30 percent each for options 2 and 3. Where the number of voters is small, there might be sufficient deviation from those patterns to tip the balance away from option 1 and toward one of the other options. But that is less likely to happen as the number of voters grows larger, since the actual *proportions* will approximate the expected ones more and more closely with an increasing number of voters. Thus, if each individual is individually more likely to vote for the "correct" option than any other, then it is likely that more individuals will vote for that "correct" option than any other — and that likelihood grows ever larger the larger the number of voters involved.

There are endless refinements and extensions of our result that might be made.[36] And there are endless further issues that jury theorems, in all their forms, must eventually confront.[37] For the purposes of this paper, we eschew these more technical issues, preferring to concentrate on the philosophical implications of 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. So long as each voter is more likely to choose the correct outcome than any other:

- The epistemically correct option is always *more likely* than any *other* option to be the plurality winner.[38] 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.[39] But at least it is always more likely to be the plurality winner than is any *other* option.
- 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.
- The correct option is more likely than any other option to be the plurality winner, *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 has a probability of more than 0.5 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 *more* quickly in the *k*-option case than it does in the 2-option case. (That should not surprise us, given that choosing the correct option with a probability of more than 0.5 is a more stringent demand in the k-option case than in the 2-option case.) Where each voter has a probability just over 0.50 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.

Table 1: **Probability that the "correct" option is the unique plurality winner**

number of options (k)	**probabilities $p_1, p_2, ..., p_k$**	**probability that option 1 (the "correct" option) is the unique plurality winner for n=**					
		11	51	101	301	601	1001
2	**0.51, 0.49**	0.527	0.557	0.580	0.636	0.688	0.737
	0.6, 0.4	0.753	0.926	0.979	♠1	♠1	♠1
3	**0.34, 0.33, 0.33**	0.268	0.338	0.358	0.407	0.449	0.489
	0.4, 0.35, 0.25	0.410	0.605	0.692	0.834	0.918	0.965
	0.5, 0.3, 0.2	0.664	0.937	0.987	♠1	♠1	♠1
4	**0.26, 0.25, 0.25, 0.24**	0.214	0.266	0.296	0.361	0.420	0.476
	0.4, 0.3, 0.2, 0.1	0.512	0.770	0.873	0.980	0.998	♠1
	0.5, 0.3, 0.1, 0.1	0.708	0.939	0.987	♠1	♠1	♠1
5	**0.21, 0.2, 0.2, 0.2, 0.19**	0.157	0.214	0.243	0.308		
	0.3, 0.2, 0.2, 0.2, 0.1	0.360	0.653	0.812	0.980		
	0.35, 0.2, 0.15, 0.15, 0,15	0.506	0.883	0.974	♠1		

[Table 1 about here]

The 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.

III. 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.[40] Contemporary social choice theorists tend to see a sharp disjunction at this point in his text[41]; and in terms of technique and methodology there certainly is.[42] 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.[43]

Condorcet's own analysis at this point of the *Essai* is notoriously opaque.[44] 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.[45] 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 escaped the notice of more philosophical commentators on epistemic democracy.[46]

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. 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* more likely to be correct in each pairwise choice (that is, 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 (that is, 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 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.[48]

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.

Appendix 2 reports the probability that the correct outcome emerges 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 restrict our attention to cases where $k \leq 3$ and to cases where the size of the electorate is 71 or smaller. Appendix 2 (Table 4) reports the probabilities of the correct option (and each of the incorrect ones) emerging as the winner under those various decision rules for 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. For an even simpler presentation, we here report — in Table 2 — just one of the cases represented in that larger Appendix 2 table, the case where there are 51 voters.

[Table 2 about here]

Table 2: **Probability that the "correct" option is the unique winner for n=51**

	number of options k=	probabilities $p_1, p_2, ..., p_k$	probability that option 1 (the "correct" option) is the unique winner among n= 51 voters under the following decision rules:				
			plurality	pairwise Condorcet	Borda	Hare	Coombs
Scenario 1	**2**	**0.51, 0.49**	0.557	0.557	0.557	0.557	0.557
Scenario 2	**3**	**0.60, 0.30, 0.10**	0.988	0.993	0.995	0.993	0.993
Scenario 3	**3**	**0.51, 0.25, 0.24**	0.972	0.991	0.994	0.989	0.993
Scenario 4	**3**	**0.40, 0.30, 0.30**	0.666	0.740	0.760	0.737	0.775
Scenario 5	**3**	**0.34, 0.33, 0.33**	0.333	0.348	0.360	0.369	0.372
Scenario 6	**3**	**0.335, 0.3325, 0.3325**	0.311	0.315	0.326	0.338	0.339

Definitions[49]:

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 m based on the candidate's ranking in a voter's preference ordering; that is, the candidate ranked first receives m points, the second one m-l, .., 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."

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*.[50]

As we have noticed, 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. 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. That has already been noted 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.

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 below $1/k$, all of these decision rules will 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 Scenario 5 and especially Scenario 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.[51] If Young's result (based though it is on rather different assumptions) can be applied, then there is every reason to believe that the epistemic strength of the Condorcet or Borda rules, anyway, will even exceed that of the plurality rule reported in Table 1.

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.[52] 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 voters, 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).

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 their varying proceduralist merits, pretty much without fear of any epistemic consequences.[53]

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.

An interesting consequence of the same mechanism is that, as the number of voters grows large, the risk of "cycling" over options declines toward the vanishing point. Were voters equally likely to support every option as every other, the opposite would occur. But just assuming voters are slightly more likely to support "correct" option than any other, the risk of cycling disappears, as is suggested by Table 2 and shown more formally in Appendix 3.

IV. Conclusions

Social choice theorists and electoral reformers debate 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.[54] 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. 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.[55] Assuming there are any truths to be found through politics, democracy has great epistemic merits, in any of its many forms.

Appendix 1: A Simple *k*-option Condorcet Jury Model

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

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

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

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

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 (where $i \neq j$), 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

$$P_i := P(\textit{for all } j \neq i,\ X_i > X_j) = P(X_1=n_1, X_2=n_2, \ldots, X_k=n_k : \langle n_1, n_2, \ldots, n_k \rangle \in N_i)$$

$$= \sum_{\langle n_1, n_2, \ldots, n_k \rangle \in N_i} \frac{n!}{n_1!\, n_2! \ldots n_k!} p_1^{n_1} p_2^{n_2} \ldots p_k^{n_k},$$

where $N_i := \{\langle n_1, n_2, \ldots, n_k \rangle : (\textit{for all } j,\ n_j \geq 0)\ \&\ (\sum_j n_j = n)\ \&\ (\textit{for all } j \neq i,\ n_i > n_j)\}$ (set of all k-tuples of votes for the k options for which option i is the plurality winner). Moreover, if, for all $j \neq i$, $p_i > p_j$, then, for all $j \neq i$, $P_i > P_j$.

Proof. The formula for P_i follows immediately from the above stated probability function for the joint distribution of X_1, X_2, ..., X_k. We will now prove that if, for all $j \neq i$, $p_i > p_j$, then, for all $j \neq i$, $P_i > P_j$.

Suppose $j \neq i$. First note that, for any k-tuple of non-negative integers $\langle n_1, n_2, ..., n_k \rangle$, we have $\langle n_1, n_2, ..., n_k \rangle \in N_j$ if and only if $\langle n'_1, n'_2, ..., n'_k \rangle \in N_i$, where $n'_i = n_j$, $n'_j = n_i$ and, for all $h \notin \{i, j\}$, $n_h = n'_h$.

Then

$$P_j = \sum_{\langle n_1, n_2, ..., n_k \rangle \in N_j} \frac{n!}{n_1!\, n_2! \ldots n_k!} p_i^{n_i} p_j^{n_j} \prod_{h \notin \{i, j\}} p_h^{n_h}$$

(notational variant of the formula for P_j)

$$= \sum_{\langle n'_1, n'_2, ..., n'_k \rangle \in N_i} \frac{n!}{n'_1!\, n'_2! \ldots n'_k!} p_i^{n'_j} p_j^{n'_i} \prod_{h \notin \{i, j\}} p_h^{n'_h}$$

(since, as noted above, $\langle n_1, n_2, ..., n_k \rangle \in N_j$ if and only if $\langle n'_1, n'_2, ..., n'_k \rangle \in N_i$, where $n'_i = n_j$, $n'_j = n_i$ and, for all $h \notin \{i, j\}$, $n_h = n'_h$).

Also,

$$P_i = \sum_{\langle n'_1, n'_2, ..., n'_k \rangle \in N_i} \frac{n!}{n'_1!\, n'_2! \ldots n'_k!} p_i^{n'_i} p_j^{n'_j} \prod_{h \notin \{i, j\}} p_h^{n'_h}$$

(notational variant of the formula for P_i).

Now, for every $\langle n'_1, n'_2, ..., n'_k \rangle \in N_i$, we have $n'_i > n'_j$, and therefore, if $p_i > p_j$, then $p_i^{n'_i} p_j^{n'_j} > p_j^{n'_i} p_i^{n'_j}$, whence, for every $\langle n'_1, n'_2, ..., n'_k \rangle \in N_i$,

$$\frac{n!}{n'_1!\, n'_2! \ldots n'_k!} p_i^{n'_i} p_j^{n'_j} \prod_{h \notin \{i, j\}} p_h^{n'_h} > \frac{n!}{n'_1!\, n'_2! \ldots n'_k!} p_i^{n'_j} p_j^{n'_i} \prod_{h \notin \{i, j\}} p_h^{n'_h},$$

and therefore

$$\sum_{\langle n'_1, n'_2, \ldots, n'_k\rangle \in N_i} \frac{n!}{n'_1!\, n'_2! \ldots n'_k!} p_i^{n'_i} p_j^{n'_j} \prod_{h \notin \{i,j\}} p_h^{n'_h}$$

$$> \sum_{\langle n'_1, n'_2, \ldots, n'_k\rangle \in N_i} \frac{n!}{n'_1!\, n'_2! \ldots n'_k!} p_i^{n'_j} p_j^{n'_i} \prod_{h \notin \{i,j\}} p_h^{n'_h},$$

and thus $P_i > P_j$, as required. **Q.E.D.**

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.

$P(\textit{for all } j \neq i,\ X_i > X_j) \to 1$ as $n \to \infty$.

Sketch proof. Consider the vector of random variables $\underline{X}^* = \langle X^*_1, X^*_2, \ldots, X^*_k\rangle$, where, for each i, $X^*_i := X_i / n$. The joint distribution of the X^*_i is a multinomial distribution with mean vector $\underline{p} = \langle p_1, p_2, \ldots, p_k\rangle$ 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 $s_{ij} = -p_i p_j$ if $i \neq j$. By the central limit theorem, for large n, $(\underline{X}^* - \underline{p})\sqrt{(n)}$ has an approximate multivariate normal distribution $N(\underline{0}, \Sigma)$. This implies that, for large n, $\underline{X}^* - \underline{p}$ has an approximate multivariate normal distribution $N(\underline{0}, {}^1/_n \Sigma)$. Let $f_n : \mathbf{R}^k \to \mathbf{R}$ be the corresponding density function for $\underline{X}^* - \underline{p}$. Using this density function, the probability that option x_i will win under plurality voting is given by $P(\textit{for all } j \neq i,\ X^*_i > X^*_j) \approx \int_{\underline{t} \in W_i} f_n(\underline{t}) d\underline{t}$, where $W_i := \{\underline{t} = \langle t_1, t_2, \ldots, t_k\rangle \in \mathbf{R}^k : \text{for all } j \neq i,\ p_i + t_i > p_j + t_j\}$. Since, by assumption, for all $j \neq i$, $p_i > p_j$, there exists an $\varepsilon > 0$ such that $S_{\underline{0},\varepsilon} \subseteq W_i$, where $S_{\underline{0},\varepsilon}$ is a sphere around $\underline{0}$ with radius ε. Then, since f_n is nonnegative, $\int_{\underline{t} \in W_i} f_n(\underline{t}) d\underline{t} \geq \int_{\underline{t} \in S_{\underline{0},\varepsilon}} f_n(\underline{t}) d\underline{t}$. But, as f_n is the density function corresponding to $N(\underline{0}, {}^1/_n \Sigma)$, $\int_{\underline{t} \in S_{\underline{0},\varepsilon}} f_n(\underline{t}) d\underline{t} \to 1$ as $n \to \infty$, and the desired result follows. **Q.E.D.**

<u>Appendix 2:</u> 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, ..., $P_{k!}$, over the k options, x_1, x_2, ..., x_k, we will assume that each voter/juror has independent probabilities p^*_1, p^*_2, ..., $p^*_{k!}$ of submitting P_1, P_2, ..., $P_{k!}$ as his/her preference ordering, respectively (where $\sum_i p_i = 1$).

Now let X^*_1, X^*_2, ..., $X^*_{k!}$ be the random variables whose values are the numbers of voters/jurors submitting the orderings P_1, P_2, ..., $P_{k!}$, respectively.

Again, the joint distribution of X^*_1, X^*_2, ..., $X^*_{k!}$ is a multinomial distribution with the following probability function:

$$P(X^*_1=n_1, X^*_2=n_2, ..., X^*_{k!}=n_{k!}) = \frac{n!}{n_1!\, n_2! \ldots n_{k!}!}\, {p^*_1}^{n_1} {p^*_2}^{n_2} \ldots {p^*_{k!}}^{n_{k!}}.$$

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

$$\frac{p_{i_1}}{1} * \frac{p_{i_2}}{1-p_{i_1}} * \frac{p_{i_3}}{1-(p_{i_1}+p_{i_2})} * \ldots * \frac{p_{i_{k-1}}}{1-(p_{i_1}+p_{i_2}+\ldots+p_{i_{k-2}})} * 1.$$

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**

	Scenario 2:	**Scenario 3:**	**Scenario 4:**	**Scenario 5:**	**Scenario 6:**
	$k = 3$	$k = 3$	$k=3$	$k=3$	$k=3$
p **option** x_1 **(correct)**	0.6	0.51	0.40	0.34	0.335
p **option** x_2	0.3	0.25	0.30	0.33	0.3325
p **option** x_3	0.1	0.24	0.30	0.33	0.3325
$p\ x_1 > x_2 > x_3$	0.450	0.260	0.200	0.170	0.1675
$p\ x_1 > x_3 > x_2$	0.150	0.250	0.200	0.170	0.1675
$p\ x_2 > x_1 > x_3$	0.257	0.170	0.171	0.168	0.1669
$p\ x_2 > x_3 > x_1$	0.043	0.080	0.129	0.162	0.1656
$p\ x_3 > x_1 > x_2$	0.067	0.161	0.171	0.168	0.1669
$p\ x_3 > x_2 > x_1$	0.033	0.079	0.129	0.162	0.1656

[Table 3 about here]

Based on these frequencies of various preference orderings, we can then calculate the probability of each of the options emerging as the winner under each of the decision rules, under each of the scenarios under consideration. Those probabilities are reported in Table 4. By stipuation, option 1 is the "correct" outcome and options 2 and 3 incorrect ones. The probabilities of the correct option 1 emerging as the winner under each rule and each scenario appears in bold type. Where the probabilities of all the options winning do not sum to one, that is because the decision rule is sometimes indecisive, yielding no winner.

[Table 4 about here]

Table 4: Probability that the "correct" option is the unique winner under various scenarios

	Scenario 1: k **= 2** p_1 **= 0.51** p_2 = 0.49 **probability of victory by:**		**Scenario 2:** k **= 3** p_1 **= 0.6** p_2 = 0.3 p_3 = 0.1 **probability of victory by:**			**Scenario 3:** k **= 3** p_1 **= 0.51** p_2 = 0.25 p_3 = 0.24 **probability of victory by:**		
	option 1	option 2	**option 1**	option 2	option 3	**option 1**	option 2	option 3
plurality								
n **= 11**	**0.527**	0.473	**0.821**	0.106	0.002	**0.699**	0.083	0.073
n **= 31**	**0.545**	0.455	**0.957**	0.028	0.000	**0.915**	0.031	0.025
n **= 51**	**0.557**	0.443	**0.988**	0.008	0.000	**0.972**	0.011	0.008
n **= 71**	**0.567**	0.433	**0.996**	0.002	0.000	**0.991**	0.004	0.002
pairwise Condorcet								
n **= 11**	**0.527**	0.473	**0.877**	0.121	0.001	**0.805**	0.090	0.077
n **= 31**	**0.545**	0.455	**0.973**	0.027	0.000	**0.959**	0.020	0.015
n **= 51**	**0.557**	0.443	**0.993**	0.007	0.000	**0.991**	0.005	0.003
n **= 71**	**0.567**	0.433	**0.998**	0.002	0.000	**0.998**	0.001	0.001
Borda								
n **= 11**	**0.527**	0.473	**0.864**	0.090	0.001	**0.805**	0.073	0.062
n **= 31**	**0.545**	0.455	**0.976**	0.017	0.000	**0.967**	0.014	0.010
n **= 51**	**0.557**	0.443	**0.995**	0.004	0.000	**0.994**	0.003	0.002
n **= 71**	**0.567**	0.433	**0.999**	0.001	0.000	**0.999**	0.001	0.000
Hare								
n **= 11**	**0.527**	0.473	**0.875**	0.123	0.002	**0.801**	0.106	0.093
n **= 31**	**0.545**	0.455	**0.973**	0.027	0.000	**0.955**	0.025	0.020
n **= 51**	**0.557**	0.443	**0.993**	0.007	0.000	**0.989**	0.006	0.004
n **= 71**	**0.567**	0.433	**0.998**	0.002	0.000	**0.997**	0.002	0.001

Coombs								
***n* = 11**	**0.527**	0.473	**0.877**	0.122	0.001	**0.816**	0.099	0.085
***n* = 31**	**0.545**	0.455	**0.973**	0.027	0.000	**0.962**	0.022	0.016
***n* = 51**	**0.557**	0.443	**0.993**	0.007	0.000	**0.991**	0.005	0.003
***n* = 71**	**0.567**	0.433	**0.998**	0.002	0.000	**0.998**	0.001	0.001

	Scenario 4: k **= 2** p_1 **= 0.40** p_2 = 0.30 p_3 = 0.30 **probability of victory by:**			**Scenario 5:** k **= 3** p_1 **= 0.34** p_2 = 0.33 p_3 = 0.33 **probability of victory by:**			**Scenario 6:** k **= 3** p_1 **= 0.335** p_2 = 0.3325 p_3 = 0.3325 **probability of victory by:**		
	option 1	option 2	option 3	**option 1**	option 2	option 3	**option 1**	option 2	option 3
plurality									
n **= 11**	**0.419**	0.178	0.178	**0.268**	0.245	0.245	**0.256**	0.251	0.251
n **= 31**	**0.581**	0.162	0.162	**0.322**	0.282	0.282	**0.302**	0.292	0.292
n **= 51**	**0.666**	0.133	0.133	**0.338**	0.286	0.286	**0.311**	0.299	0.299
n **= 71**	**0.726**	0.108	0.108	**0.344**	0.283	0.283	**0.313**	0.298	0.298
pairwise Condorcet									
n **= 11**	**0.514**	0.209	0.209	**0.327**	0.297	0.297	**0.312**	0.304	0.304
n **= 31**	**0.652**	0.146	0.146	**0.339**	0.288	0.288	**0.313**	0.301	0.301
n **= 51**	**0.740**	0.109	0.109	**0.348**	0.283	0.283	**0.315**	0.299	0.299
n **= 71**	**0.802**	0.083	0.083	**0.356**	0.279	0.279	**0.317**	0.298	0.298
Borda									
n **= 11**	**0.505**	0.195	0.195	**0.316**	0.286	0.286	**0.301**	0.293	0.293
n **= 31**	**0.666**	0.140	0.140	**0.344**	0.292	0.292	**0.318**	0.305	0.305
n **= 51**	**0.760**	0.103	0.103	**0.360**	0.292	0.292	**0.326**	0.309	0.309
n **= 71**	**0.823**	0.076	0.076	**0.371**	0.291	0.291	**0.330**	0.311	0.311
Hare									
n **= 11**	**0.527**	0.236	0.236	**0.352**	0.324	0.324	**0.338**	0.331	0.331
n **= 31**	**0.655**	0.172	0.172	**0.365**	0.318	0.318	**0.341**	0.329	0.329
n **= 51**	**0.737**	0.126	0.126	**0.369**	0.307	0.307	**0.338**	0.323	0.323
n **= 71**	**0.800**	0.100	0.100	**0.382**	0.309	0.309	**0.345**	0.327	0.327
Coombs									
n **= 11**	**0.539**	0.231	0.231	**0.354**	0.323	0.323	**0.338**	0.331	0.331
n **= 31**	**0.675**	0.162	0.162	**0.368**	0.316	0.316	**0.342**	0.329	0.329
n **= 51**	**0.755**	0.119	0.119	**0.372**	0.306	0.306	**0.339**	0.323	0.323
n **= 71**	**0.815**	0.092	0.092	**0.385**	0.307	0.307	**0.346**	0.327	0.327

The difference between 1 and the sum of the probabilities of victory by the different options is the probability that the procedure will be indecisive.

Appendix 3: An Implication of the *k*-option Condorcet Jury Mechanism for the Probability of Cycles[56]

Although the pairwise Condorcet winner criterion may seem an attractive democratic decision procedure, it is famously threatened by Condorcet's paradox: pairwise majority voting may lead to cyclical collective preferences. But how probable is the occurrence of cycles?

An important body of literature addressing this question uses the so-called 'impartial culture' assumption[57]. Given the $k!$ logically possible strict preference orderings, $P_1, P_2, ..., P_{k!}$, over k options, $x_1, x_2, ..., x_k$, it is assumed that all of these orderings are equally likely to be submitted by an individual voter/juror, i.e. each voter/juror has independent probabilities $p^*_1 = p^*_2 = ... = p^*_{k!} = 1/k!$ of submitting $P_1, P_2, ..., P_{k!}$ as his/her preference ordering, respectively. Given this perfect equiprobability assumption, the probability of the existence of a Condorcet winner *decreases* with increases in the number of voters/jurors as well as with increases in the number of options. The larger the electorate, the harder it would seem to generate a Condorcet winning outcome.

This theoretical result is strikingly at odds with our empirical observations. Cycles are much less common in the real world than some of the social-choice-theoretic literature would lead us to expect[58]. But the result also seems hard to reconcile with one of the main points of this paper. Given suitable minimal assumptions about the competence of individual voters/jurors, the point has been that several plausible social choice procedures produce the 'correct' option as their unique winning outcome with a probality *increasing in the number of voters/jurors*. In particular, under these assumptions the 'correct' option is also increasingly likely to emerge as the unique pairwise Condorcet winner as the number of voters/jurors increases.

The response to this apparent clash of theoretical results is that the present assumptions about voter/juror competence break the 'impartial culture' assumption. Crucially, the assumption that each voter is more likely, however slightly, to choose the 'correct' option than any other is a violation of the assumption that all logically possible preference orderings are equally probable to occur. This raises the question of how much deviation from this equiprobability assumption is necessary to avoid the standard result on the probability of cycles. As we will see in this final appendix, the very same mechanism that underlies the *k*-option Condorcet jury theorem has an implication for this question too.

Using the three-option case as a simple illustration, we will now show that the impartial culture assumption can be seen as an extreme limiting case the slightest systematic deviation from which is already sufficient to circumvent the standard cycling result, provided the electorate is sufficiently large[59].

Suppose there are n voters/jurors (n odd) and three options, x, y, z. For simplicity, we will only consider strict preference orderings. There are 6 logically possible such orderings of the options:

label	P_{X1}	P_{Y2}	P_{Z1}	P_{X2}	P_{Y1}	P_{Z2}
1st	z	z	y	y	x	x
2nd	x	y	z	x	y	z
3rd	y	x	x	z	z	y

Let $n(P_{X1})$, $n(P_{X2})$, $n(P_{Y1})$, $n(P_{Y2})$, $n(P_{Z1})$, $n(P_{Z2})$ be the numbers of voters/jurors submitting orderings P_{X1}, P_{X2}, P_{Y1}, P_{Y2}, P_{Z1}, P_{Z2}, respectively. The vector $<n(P_{X1}), n(P_{X2}), n(P_{Y1}), n(P_{Y2}), n(P_{Z1}), n(P_{Z2})>$ is called an *anonymous preference profile*.

Proposition 3. (Nicholas Miller) The anonymous profile $<n(P_{X1}), n(P_{X2}), n(P_{Y1}), n(P_{Y2}), n(P_{Z1}), n(P_{Z2})>$ generates a cycle under pairwise majority voting if and only if

$$\begin{aligned} &[\,[n(P_{X1}) > n(P_{X2}) \;\&\; n(P_{Y1}) > n(P_{Y2}) \;\&\; n(P_{Z1}) > n(P_{Z2})] \\ &\quad or\ [n(P_{X1}) < n(P_{X2}) \;\&\; n(P_{Y1}) < n(P_{Y2}) \;\&\; n(P_{Z1}) < n(P_{Z2})]\,] \\ &\&\ |n(P_{X1}) - n(P_{X2})| < n'/2 \\ &\&\ |n(P_{Y1}) - n(P_{Y2})| < n'/2 \\ &\&\ |n(P_{Z1}) - n(P_{Z2})| < n'/2, \end{aligned}$$

where $n' = |n(P_{X1}) - n(P_{X2})| + |n(P_{Y1}) - n(P_{Y2})| + |n(P_{Z1}) - n(P_{Z2})|$.

Now let $p_{X1}, p_{X2}, p_{Y1}, p_{Y2}, p_{Z1}, p_{Z2}$ be the probabilities that an individual voter/juror submits the orderings $P_{X1}, P_{X2}, P_{Y1}, P_{Y2}, P_{Z1}, P_{Z2}$, respectively (where the sum of the probabilities is 1). An *impartial culture* is the situation in which $p_{X1} = p_{X2} = p_{Y1} = p_{Y2} = p_{Z1} = p_{Z2}$.

Let $X_{X1}, X_{X2}, X_{Y1}, X_{Y2}, X_{Z1}, X_{Z2}$ be the random variables whose values are the numbers of voters/jurors with orderings $P_{X1}, P_{X2}, P_{Y1}, P_{Y2}, P_{Z1}, P_{Z2}$, respectively.

As in the model of appendix 2, the joint distribution of $X_{X1}, X_{X2}, X_{Y1}, X_{Y2}, X_{Z1}, X_{Z2}$ is a multinomial distribution with the following probability function:

$$P(X_{X1}=n_{X1}, X_{X2}=n_{X2}, \ldots, X_{Z2}=n_{Z2}) = \frac{n!}{n_{X1}!\, n_{X2}! \ldots n_{Z2}!} p_{X1}^{n_{X1}} p_{X2}^{n_{X2}} \ldots p_{Z2}^{n_{Z2}}.$$

Proposition 4. Suppose

$$
\begin{aligned}
&[\,[p_{X1} < p_{X2} \text{ or } p_{Y1} < p_{Y2} \text{ or } p_{Z1} < p_{Z2}] \\
&\;\&\; [p_{X1} > p_{X2} \text{ or } p_{Y1} > p_{Y2} \text{ or } p_{Z1} > p_{Z2}]\,] \\
&\text{or } |p_{X1} - p_{X2}| > n'/2 \\
&\text{or } |p_{Y1} - p_{Y2}| > n'/2 \\
&\text{or } |p_{Z1} - p_{Z2}| > n'/2,
\end{aligned}
$$

where $n' = |p_{X1} - p_{X2}| + |p_{Y1} - p_{Y2}| + |p_{Z1} - p_{Z2}|$. Then the probability that there will be no cycle under pairwise majority voting tends to 1 as n tends to infinity.

Sketch proof. Consider the vector of random variables $\underline{X}^* = \langle X^*_{X1}, X^*_{X2}, X^*_{Y1}, X^*_{Y2}, X^*_{Z1}, X^*_{Z2}\rangle$, where, for each $i \in \{X1, X2, Y1, Y2, Z1, Z2\}$, $X^*_i := X_i/n$. The joint distribution of the X^*_i is a multinomial distribution with mean vector $\underline{p} = \langle p_{X1}, p_{X2}, p_{Y1}, p_{Y2}, p_{Z1}, p_{Z2}\rangle$ and with variance-covariance matrix $\Sigma = (s_{ij})$, where, for each $i, j \in \{X1, X2, Y1, Y2, Z1, Z2\}$, $s_{ij} = p_i(1-p_i)$ if $i=j$ and $s_{ij} = -p_ip_j$ if $i \neq j$. Again by the central limit theorem, for large n, $(\underline{X}^*-\underline{p})\sqrt{(n)}$ has an approximate multivariate normal distribution $N(\underline{0}, \Sigma)$, and $\underline{X}^*-\underline{p}$ has an approximate multivariate normal distribution $N(\underline{0}, {}^1/_n \Sigma)$. Let $f_n : \mathbf{R}^6 \to \mathbf{R}$ be the corresponding density function for $\underline{X}^*-\underline{p}$. From proposition **3** we can infer, using this density function, that the probability that there will be no cycle under majority voting is given by $\int_{\underline{t} \in W} f_n(\underline{t})d\underline{t}$, where

$$
\begin{aligned}
W := \{&\underline{t} = \langle t_{X1}, t_{X2}, t_{Y1}, t_{Y2}, t_{Z1}, t_{Z2}\rangle \in \mathbf{R}^6 : \\
&[\,[\,p_{X1}+t_{X1} < p_{X2}+t_{X2} \text{ or } p_{Y1}+t_{Y1} < p_{Y2}+t_{Y2} \text{ or } p_{Z1}+t_{Z1} < p_{Z2}+t_{Z2}] \\
&\;\&\; [p_{X1}+t_{X1} > p_{X2}+t_{X2} \text{ or } p_{Y1}+t_{Y1} > p_{Y2}+t_{Y2} \text{ or } p_{Z1}+t_{Z1} > p_{Z2}+t_{Z2}]\,] \\
&\text{or } |(p_{X1}+t_{X1}) - (p_{X2}+t_{X2})| > n'/2 \\
&\text{or } |(p_{Y1}+t_{Y1}) - (p_{Y2}+t_{Y2})| > n'/2 \\
&\text{or } |(p_{Z1}+t_{Z1}) - (p_{Z2}+t_{Z2})| > n'/2,
\end{aligned}
$$

where $n' = |(p_{X1}+t_{X1}) - (p_{X2}+t_{X2})| + |(p_{Y1}+t_{Y1}) - (p_{Y2}+t_{Y2})| + |p_{Z1}+t_{Z1}) - (p_{Z2}+t_{Z2})|\}$.

Note that, by assumption, $\underline{0} \in W$, and since all relevant inequalities satisfied by p_{X1}, p_{X2}, p_{Y1}, p_{Y2}, p_{Z1}, p_{Z2} are strict, there exists an $\varepsilon > 0$ such that $S_{\underline{0},\varepsilon} \subseteq W$, where $S_{\underline{0},\varepsilon}$ is a sphere around $\underline{0}$ with radius ε. Then, since f_n is nonnegative, $\int_{\underline{t} \in W} f_n(\underline{t}) d\underline{t} \geq \int_{\underline{t} \in S_{\underline{0},\varepsilon}} f_n(\underline{t}) d\underline{t}$. But, as f_n is the density function corresponding to $N(\underline{0}, {}^1/_n \Sigma)$, $\int_{\underline{t} \in S_{\underline{0},\varepsilon}} f_n(\underline{t}) d\underline{t} \to 1$ as $n \to \infty$, and hence $\int_{\underline{t} \in W} f_n(\underline{t}) d\underline{t} \to 1$ as $n \to \infty$, as required. **Q.E.D.**

Note that the condition of proposition **4** is already satisfied if at least one of $p_{X1} < p_{X2}$, $p_{Y1} < p_{Y2}$, $p_{Z1} < p_{Z2}$ and at least one of $p_{X1} > p_{X2}$, $p_{Y1} > p_{Y2}$, $p_{Z1} > p_{Z2}$ are satisfied. For instance, the condition is satisfied if $p_{X1} = 1/6 - \varepsilon$, $p_{Y1} = 1/6 + \varepsilon$ and $p_{X2} = p_{Y2} = p_{Z1} = p_{Z2} = 1/6$.

Proposition **4** implies that, given suitable systematic, however slight, deviations from an impartial culture, the probability that there will be a cycle under pairwise majority voting vanishes as the size of the electorate increases.

The mechanism underlying this result is formally similar to the mechanism underlying the k-option Condorcet jury theorem. If p_{X1}, p_{X2}, p_{Y1}, p_{Y2}, p_{Z1}, p_{Z2} are the probabilities that an individual voter/juror submits the orderings P_{X1}, P_{X2}, P_{Y1}, P_{Y2}, P_{Z1}, P_{Z2}, respectively, then np_{X1}, np_{X2}, np_{Y1}, np_{Y2}, np_{Z1}, np_{Z2} are the expected frequencies of these orderings amongst the n orderings submitted by an electorate of n voters/jurors. If n is small, the actual frequencies may differ substantially from this pattern, but, as n increases, the actual frequencies approximate the expected frequencies increasingly closely in relative terms, by the law of large numbers. In particular, provided the probabilities satisfy the condition of proposition **4**, the actual anonymous profile $<n(P_{X1}), n(P_{X2}), n(P_{Y1}),$

$n(P_{Y2})$, $n(P_{Z1})$, $n(P_{Z2})$> is thus increasingly likely to satisfy the negation of the condition of proposition **3** and hence decreasingly likely to generate a cycle.

However, if the probabilities deviate systematically from an impartial culture so as to replicate the pattern of Condorcet's paradox, the probability that there will be a cycle under pairwise majority voting tends to 1 as n tends to infinity.

Proposition 5. Suppose

$$\begin{aligned} &[\,[p_{X1} > p_{X2} \,\&\, p_{Y1} > p_{Y2} \,\&\, p_{Z1} > p_{Z2}] \\ &\quad or\ [p_{X1} < p_{X2} \,\&\, p_{Y1} < p_{Y2} \,\&\, p_{Z1} < p_{Z2}]\,] \\ \&\ \ &|p_{X1} - p_{X2}| < n'/2 \\ \&\ \ &|p_{Y1} - p_{Y2}| < n'/2 \\ \&\ \ &|p_{Z1} - p_{Z2}| < n'/2, \end{aligned}$$

where $n' = |p_{X1} - p_{X2}| + |p_{Y1} - p_{Y2}| + |p_{Z1} - p_{Z2}|$. Then the probability that there will be a cycle under pairwise majority voting tends to 1 as n tends to infinity.

Proof analogous to the proof of proposition **4**.

An impartial culture is a rather unstable limiting case. We have seen that in any ε-neighbourhood of an impartial culture there are situations, as described by proposition **4**, in which the probability of the occurrence of a cycle converges to 0 as the size of the electorate increases. Likewise, there are situations, as described by proposition **5**, in which the probability of the occurrence of a cycle converges to 1 as the size of the electorate increases. It is an empirical question which of these situations is the more common one. Logically, the 'mostly disjunctive' condition of proposition **4** is less demanding than the 'mostly conjunctive' condition of proposition **5**. Moreover, given the lack of empirical evidence of

cycles, the condition of proposition **4** is arguably the empirically more plausible one.

References

Arrow, Kenneth J. 1963. Social Choice and Individual Values. 2nd edn. New Haven, Conn.: Yale University Press.

Austen-Smith, David and Jeffrey S. Banks. 1996. Information aggregation, rationality and the Concordet jury theorem. American Political Science Review, 90 (#1: March): 34-45.

Barry, Brian. 1964. The public interest. Proceedings of the Aristotelian Society, Supplement, 38: 1-18.

Barry, Brian. 1965. Political Argument. London: Routledge & Kegan Paul.

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 andWelfare, 6 (#2): 157-65.

Black, Duncan. 1958. The Theory of Committees and Elections. Cambridge: Cambridge University Press.

Borda, Jean-Charles de. 1784. Mémoire sur les élections au scrutin par M. de Borda. Pp. 657-65 Mémoires de l'Académie Royale des Sciences année 1781. Paris: l'Imprimerie Royale. Translated and reprinted in: Sommerlad and McLean 1988, pp. 122-8; McLean and Hewitt 1994, pp. 114-9; McLean and Urken 1995, pp. 83-9; Sommerlad and McLean 1988, pp. 122-8.

Borland, Philip J. 1989. Majority systems and the Condorcet jury theorem. Statistician 38 (#3): 181-9.

Brennan, H. Geoffrey. 1995. Condorcet [review of McLean and Hewitt 1994]. History of Economics Review #23 (Winter): 157-9.

Coleman, Jules and John Ferejohn. 1986. Democracy and social choice. Ethics, 97 (#1: Oct): 6-25.

Cohen, Joshua. 1986. An epistemic conception of democracy. Ethics, 97 (1: Oct): 26-38.

Coleman, Jules L. 1989. Rationality and the justification of democracy. Pp. 194-220 in Politics and Process: New Essays in Democratic Thought, ed. Geoffrey Brennan and Loren E. Lomasky. Cambridge: Cambridge University Press.

Condorcet, Marie Jean Antoine Nicolas de Caritat, Marquis de. 1784. Sur les élections par scrutin. Pp. 31-4 in Histoire de l'Académie Royale des Sciences année 1781. Paris: De l'imprimerie royale. Translated and reprinted in: Sommerlad and McLean 1988, pp. 119-22; McLean and Hewitt 1994, pp. 111-3; McLean and Urken 1995, pp. 81-3.

Condorcet, Marie Jean Antoine Nicolas de Caritat, Marquis de. 1785. Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. Paris: l'Imprimerie Royale [fascimile edition New York: Chelsea, 1972]. Comprised of: Discours préliminaire, pp i-cxci; Essai, pp. 1-304. Translated and excerpted in: Sommerlad and McLean 1989, pp. 69-80 (corresponding to Condorcet 1785/1972, pp. lvi-lxx); Sommerlad and McLean 1989, pp. 81-9 (corresponding to Condorcet 1785/1972, pp. clxviii-clxxix); Sommerlad and McLean 1989, pp. 90-108 (corresponding to Condorcet 1785/1972, pp. 119-36); Sommerlad and McLean 1989, pp. 109-18 (corresponding to Condorcet 1785/1972, pp. 287-96); McLean and Hewitt 1994, pp. 120-30 (corresponding to Condorcet 1785/1972, pp. lvi-lxx); McLean and Hewitt 1994, pp. 130-8 (corresponding to Condorcet 1785/1972, pp. clxviii-clxxix) ; McLean and Urken 1995, pp. 91-113 (corresponding to Condorcet 1785/1972, pp. 279-304).

Condorcet, Marie Jean Antoine Nicolas de Caritat, Marquis de. 1788. On the constitution and the functions of provincial assemblies. Translated and

excerpted in: McLean and Hewitt 1994, pp. 139-68; McLean and Urken 1995, pp. 113-44; Sommerlad and McLean 1989, pp. 129-66.

Condorcet, Marie Jean Antoine Nicolas de Caritat, Marquis de. 1792. A survey of the principles underlying the Draft Constitution. Translated and excerpted in: McLean and Hewitt 1994, pp. 190-227; McLean and Urken 1995, pp. 145-50; Sommerlad and McLean 1989, pp. 216-22..

Coughlan, Peter J. 2000. In defense of unanimous jury verdicts: mistrials, communication and strategic voting. American Political Science Review, 94 (#2: June): 375-94.

Dahl, Robert A. 1979. Procedural democracy. Pp. 97-133 in Philosophy, Politics & Society, 5th series, ed. P. Laslett & J. Fishkin (Oxford, Blackwell.

Dummett, Michael. 1985. Voting Procedures. Oxford: Clarendon Press.

Dummett, Michael. 1997. Principles of Electoral Reform. Oxford: Clarendon Press.

Estlund, David M. 1990. Democracy without preference. Philosophical Review 49: 397-424.

Estlund, David. 1993. Making truth safe for democracy. Pp. 71-100 in The Idea of Democracy, ed David Copp, Jean Hampton and John E. Roemer. New York: Cambridge University Press.

Estlund, David. 1994. Opinion leaders, independence and Condorcet's jury theorem. Theory & Decision, 36 (#2): 131-62.

Estlund, David. 1997. Beyond fairness and deliberation: the epistemic dimension of democratic authority. Pp. 173-204 in Deliberative Democracy, ed. James Bohman and Willian Rehg. Cambridge, Mass.: MIT Press.

Estlund, David. 1998. The insularity of the reasonable: why political liberalism must admit the truth. Ethics, 108 (Jan): 252-75.

Estlund, David; Jeremy Waldron; Bernard Grofman and Scott Feld. 1989. Controversy: democratic theory and the public interest: Condorcet and Rousseau revisited. American Political Science Review, 83 (#4: Dec.): 1317-40.

Feddersen, Timothy and Wolfgang Pesendorfer. 1998. Convicting the innocent: the inferiority of unanimous jury verdicts under strategic voting. American Political Science Review, 92: 23-36.

Fung, Archon. 1995. Extended Condorcet and experimentalist models of epistemic democracy. Mimeo, Department of Political Science, MIT.

Gaus, Gerald. 1997. Does democracy reveal the voice of the people? Four takes on Rousseau. Australasian Journal of Philosophy 75 (#2): 141-62.

Gerardi, Dino. 2000. Jury verdicts and preference diversity. American Political Science Review, 94 (#2: June): 395-406.

Gehrlein, William. 1983. Condorcet's paradox. Theory & Decision, 15: 161-97.

Granger, Giles-Gaston. 1968. Condorcet. Vol. 3, pp. 213-5 in International Encyclopedia of the Social Sciences, ed. David L. Sills. London: Collier-Macmillan.

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

Grofman, Bernard and Scott L. Feld. 1988. Rosseau's general will: a Condorcetian perspective. American Political Science Review, 82 (#2: June): 567-76.

Grofman, Bernard and Guillermo Owen, eds. 1986. Information Pooling and Group Decision Making. Westport, Conn.: JAI Press.

Guarnaschelli, Serena; Richard D. McKelvey; and Thomas R. Palfrey. 2000. An experimental study of jury decision rules. American Political Science Review, 94 (#2: June): 407-24.

Kuflik, Arthur. 1977. Majority rule procedure. Pp. 296-322 in Nomos XVIII: Due Process, ed. J. Roland Pennock and John W. Chapman. New York: New York University Press.

Lahda, Krishna K. 1992. The Condorcet jury theorem, free speech and correlated votes. American Journal of Political Science, 36 (#3: Aug): 617-34.

Levin, Jonathan and Barry Naelbuff. 1995. An introduction to vote-counting schemes. Journal of Economic Perspectives, 9 (#1: Winter): 3-26.

Manin, Bernard; Adam Przeworski; and Susan C. Stokes. 1999. Introduction. Pp. 1-26 in Democracy, Accountability and Representation, ed. A. Przeworski, S. C. Stokes and B. Manin. Cambridge: Cambridge University Press.

May, Kenneth O. 1952. A set of independent, necessary and sufficient conditions for simple majority decision. Econometrica 20: 680-84.

McLean, Iain and Fiona Hewitt, trans. and ed. 1994. Condorcet: Foundations of Social Choice and Political Theory. Aldershot, Hants.: Elgar.

McLean, Iain and Arnold B. Urken, trans and ed. 1995. Classics of Social Choice. Ann Arbor: University of Michigan Press.

Merrill, Samuel, III. 1984. A comparison of efficiency of multicandidate electoral systems. American Journal of Political Science, 28 (#1: Feb): 23-48.

Mill, James. 1823. Essay on government. Reprinted as James Mill: Political Writings, ed. Terence Ball. Cambridge: Cambridge University Press, 1992.

Miller, David. 1992. Deliberative democracy and social choice. Political Studies, 40 (#5: Special Issue): 54-67. Reprinted Pp. 54-67 in Prospects for Democracy, ed. David Held. Oxford: Blackwell.

Mueller, Dennis C. 1989. Public Choice II. Cambridge: Cambridge University Press.

Riker, William H. 1983. Liberalism against Populism. San Francisco: W. Freeman.

Rousseau, Jean-Jacques. 1762. The social contract. Pp. 164-278 in The Social Contract and Discourses, trans. G. D. H. Cole. New ed. London: Everyman/Dent, 1973.

Schumpeter, Joseph A. 1950. Capitalism, Socialism & Democracy. 3rd ed. New York: Harper & Row.

Sen, Amartya. 1982. Choice, Welfare and Measurement. Oxford: Blackwell.

Shapley, Lloyd S. and Bernard Grofman. 1984. Optimizing group judgmental accuracy in the presence of interdependencies. Public Choice, 43 (#3): 329-43.

Sommerlad, Fiona and Iain McLean, trans. and ed. 1989. The Political theory of Condorcet. Social Studies Faculty Centre Working Paper 1/89. Oxford: Faculty of Social Sciences, Oxford University.

Spitz, Elaine. 1984. Majority Rule. Chatham, N.J.: Chatham House Publishers.

Tangian, A. S. 2000. Unlikelihood of Condorcet's paradox in a large society. Social Choice & Welfare, 17: 337-65.

Waldron, Jeremy. 1999. The Dignity of Legislation. Cambridge: Cambridge University Press.

Wit, Jörgen. 1998. Rational choice and the Condorcet jury theorem. Games & Economic Behavior, 22: 364-76.

Young, H. P. 1988. Condorcet's theory of voting. American Political Science Review, 82 (#4: Dec): 1231-44.

Young, Peyton. 1995. Optimal voting rules. Journal of Economic Perspectives, 9 (#1: Winter): 51-64.

Notes

* 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 and finalized while he was a visiting scholar at Harvard and MIT. 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 for comments from Geoff Brennan, Dave Estlund, David Firth, Barbara Fried, Archon Fung, Jerry Gaus, David Gauthier, Bernie Grofman, Saul Levmore, Iain McLean, David Miller, Nick Miller, Philip Pettit, John Quiggin, Mathias Risse, Mariam Thalos, Jeremy Waldron and two anonymous referees. We are also grateful to Associate Editor John Dryzek for overseeing the refereeing of this submission.

[1] Estlund (1990; 1993; 1997; 1998) is the most assiduous contemporary advocate of that position, but he has illustrous predecessors. Rousseau arguably 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."

[3] The modern *locus classicus* is Dahl 1979. See similarly Schumpeter 1950, pt 4.

[4] Coleman and Ferejohn (1986, p. 7) define "the *proceduralist* approach to the justification of collective decision" as one which "identif[ies] a set of ideals with which any 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.

[5] For elaboration of the jury theorem, see Section II below.

[6] 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).

[7] 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).

[8] 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.

[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."

[10] Surveyed, respectively, in Mueller (1989) and Dummett (1985; 1997).

[11] Referring to previous versions of the Condorcet jury theorem, Estlund (1997, p. 189), for example, notes that "the Jury Theorem assumes there are only two alternatives." He goes on to say, "For these and other reasons, the Jury Theorem approach to the epistemic value of democratic procedures is less than trustworthy."

[12] Especially, as we shall see, when there is more information available about voters' preferences than just a single vote for each voter.

[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.

[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 nothing but the truth; and even then only if the person receiving god's revelations never mistakes god's meaning.

[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 "especially as [the number of voters] is large, ... optimal [competence-based] weights do not improve substantially on simple majority rule" assigning them equal weight.

[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] 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.

[18] Any more than we attempted to resolve all the highly contentious philosophical issues in epistemology, in talking under the epistemological heading of "the truth."

[19] "If a form of election is to be just, the voters must be able to rank *each* candidate according to his merits, compared successively to the merits of *each of the others*... " Borda goes 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 *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 *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, etc., 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 naturaly seems to tend, Dahl goes on to say, quite emphatically, that: "The *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).

[20] 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.

[21] 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).

[22] "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).

[23] 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.

[24] Condorcet 1785, p. 279 ff. Black 1958, pp. 163-5.

[25] Barry 1964 , pp. 9-14; 1965, Appendix A, pp. 292-3. Kuflik 1977, pp. 305-8. Spitz 1984, p. 206. Cohen 1986, p. 35. Grofman and Feld 1988; cf. Estlund, Waldron, Grofman and Feld 1989. Martin 1993, pp. 142-4, 370-1. Gaus 1997, pp. 149-50. Estlund 1993, pp. 92-4.

[26] Grofman and Feld 1983, pp. 268 ff. Borland 1989, p. 183.

[27] Grofman and Feld 1983, pp. 273-4. Borland 1989, p. 185-6. Ladha 1992. Estlund 1994.

[28] Austen-Smith and Banks 1996. Feddersen and Pesendorfer 1998. Wit 1998. Coughlan 2000. Gerardi 2000. Guarnaschelli, McKelvey and Palfrey 2000.

[29] 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 other voter and as in each other pairwise comparison. On the face of it, this seems problematic insofar as it seems to treat probabilities in each of a voter's pairwise choices as independent of those in each of the same voter's other pairwise choices, when the probabilities in the one case may seem to constrain and be constrained by the probabilities in all the others (especially if people's preferences satisfy certain consistency conditions).

[30] 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 $p_{mean}>0.471$, although the value goes that low only when there are merely three voters.

[31] 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."

[32] Gaus 1997, p. 150.

[33] 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.

[34]These options are taken as exogenous in our analysis. The setting of the agenda over which voters choose is obviously an important issue, but it is beyond our remit here.

[35]After posting an earlier draft of this paper on the internet, our attention was drawn to Archon Fung's (1995) sketch of a similar result.

[36] 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 s itself greater than 1/2 (Grofman, Owen and Feld 1983, pp. 268-9). Similar generalizations might be made in the k-option case.

[37] Most especially, issues of strategic voting.

[38] 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.

[39] 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.

[40] 1785, pp. 122 ff. Black 1958, pp. 168-71. McLean and Hewitt 1994, pp. 38-40.

[41] 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."

[42] "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 on the modern theory of social choice," writes Young (1988, p. 1238-9). See similarly 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] Black 1958, p. 163. Granger 1968, p 214. Young 1988, p. 1232. Brennan 1995.

[44] 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).

[45] 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.

[46] 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.

[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).

[48]Otherwise, under typical election laws, their votes will be deemed "invalid" and not count.

[49]All definitions are from Mueller 1989, pp. 112-3.

[50] That is why the probability of the "correct" option being chosen by the group is sometimes lower (in Scenario 6, especially) than the probability of of each individual voting for the "correct" option: those are cases in which the decision rule in question yielded either no winner or none uniquely. Casual inspection of sample calculations in Appendix 2, Table 4 may seem to suggest that this might also 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 (that is, result in ties) in the k=3 case only when n is a multiple of 2 or 3; that affects the case of n=51 uniquely among the values of n for which Appendix 2, Table 4 calculates these probabilities.

[51] 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.

[52] Such as the apparent "reversal" with Hare and Coombs rules, discussed in the preceding note.

[53]Estlund (1997, p. 174) supposes that "democratic legitimacy requires that the procedure is procedurally fair and can be held, in terms acceptable to all reasonable citizens, to be epistemically the best among those that are better than random." The upshot of our findings here is that the second half of that standard is substantially otoise: in large electorates, pretty much any procedure is epistemically as good as any other.

[54] Cf. Black 1958, p. 163; Miller 1992, p. 56.

[55] 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.

[56] This appendix is the work of Christian List. He is greatly indebted to Nicholas Miller, the late Jeff Banks, Christian Elsholtz, Marek Kaminski and Gerry Mackie for discussion; all remaining errors are his own. A previous version of this material was presented at a workshop on "Topics in Mathematical Models of Individual and Public Choice" at the Institute for Mathematical Behavioral Sciences at the University of California at Irvine, July 2000.

[57] Gehrlein 1983.

[58] As unpublished work by Gerry Mackie shows.

[59] A more technical and comprehensive treatment of some related results can be found in a very recent paper by Tangian (2000).