

draft as at 14 Ap 10

## **Courts of Many Minds\***

ROBERT E. GOODIN

Social & Political Theory and Philosophy Programs  
Research School of Social Sciences

Australian National University  
<goodinb@coombs.anu.edu.au>

and

KAI SPIEKERMANN

Department of Government  
London School of Economics  
<k.spiekermann@lse.ac.uk>

© 2010 Robert E. Goodin and Kai Spiekermann

\* A more technical version of this paper can be found at  
<[http://personal.lse.ac.uk/spiekerk/papers/Courts\\_of\\_Many\\_Minds\\_technical.pdf](http://personal.lse.ac.uk/spiekerk/papers/Courts_of_Many_Minds_technical.pdf)>. We are  
grateful to Adrian Vermeule for comments on an earlier draft.

---

In his important recent book, *A Constitution of Many Minds*, Cass Sunstein explores three alternative approaches to constitutional interpretation:

{1} Traditionalists insist that if members of a society have long accepted a certain practice, courts should be reluctant to disturb that practice. .... [2] Populists believe that if most people believe a certain fact or accept a certain value, judges should show a degree of humility – and respect their view in the face of reasonable doubt. .... [3] Cosmopolitans believe that if many nations, or many democratic nations, reject a practice, or accept a practice, the U.S. Supreme Court should pay respectful attention (Sunstein 2009: ix-x).

Sunstein says that all three approaches share a common premise – the "many minds argument" of his book's title. The thought is that, "if many people think something, their view is entitled to consideration and respect" (Sunstein 2009: ix-x) – not just as a matter of courtesy, but because the more of them there are, the more likely they are to be right. In that respect, "The structure of the central argument is identical in all three contexts." All three approaches rest on the same formal foundations, the Condorcet Jury Theorem (CJT) (Sunstein 2009: x, 8-10).

Insofar as the CJT is the logic underlying all three approaches to constitutional interpretation, that same Theorem should provide a formal basis for adjudicating among them. In his subsequent discussion of those issues, Sunstein himself abjures formalism in favor of more context-sensitive discussions of a more lawyerly sort. The aim of this paper is to provide more formal assessments of the epistemic power of those approaches, through

suitable elaborations and extensions of the CJT apparatus.

"Is it possible to compare and to rank the three kinds of many minds argument?" Sunstein (2009: 214) asks at the end of the book. His conclusion: "For the United States, I have suggested that traditions are likely to provide the strongest basis for constitutional law, and that international practices provide the weakest. Public convictions are in the middle."

In Section II of this paper, we discuss particular problems from a CJT point of view that plague Traditionalism, Sunstein's preferred approach. (Everything said there can also be applied with minor adjustments to the Cosmopolitan approach, which we do not here model separately.) In Section III, we discuss problems that Sunstein sees from a CJT point of view with the Populist approach, showing that they might be less problematic than he imagines.

## **I. The Condorcet Jury Theorem: Basic Set-up**

Before proceeding with any of that, however, we offer a brief reminder of the basic structure of the CJT apparatus. In its original version, the Condorcet Jury Theorem applies to a group of jurors deciding between two alternatives by a majority vote. The Theorem relies on two key assumptions. One

concerns the competence of voters: each voter is assumed to vote for the

correct alternative with probability  $p > 0.5$ . The other concerns the

independence of voters: the votes of all jurors are statistically independent of one another. Provided those assumptions are satisfied, the Condorcet Jury

Theorem tells us this: the probability of a majority of jurors being correct is greater the more jurors there are; and that probability converges to 1 as the number of jurors tends to infinity.

A great many refinements, extensions and qualifications have been made to the CJT over the years. Many of them are reassuring. Thus, for example, we do not need to assume that all jurors are of identical competence: the same result usually holds for jurors of differing competence, just so long as mean competence in the jury is  $p > 0.5$  (Groffman et al. 1983; Owen et al. 1989).<sup>1</sup> And even if there are more than just two options, essentially the same result still holds for plurality voting over any number of alternatives (List and Goodin 2001).

In respect of the key independence and competence assumptions, however, worries remain. While sometimes presented as distinct worries (Estlund 2008: 225-8), both are best seen as aspects of one and the same worry – which is that there might be some common influences that systematically affect many voters at once, thus violating the independence assumption. (All the CJT requires by way of competence, recall, is that voters be "better than random": how could they be worse than random, except in that way?) Such violations of the independence assumption are largely what lead Sunstein (2009: 171-3) to reject Populism. Hence the importance of our

<sup>1</sup> 'Usually' because you can get extreme distributions of competences for which it is not true: e.g. if three voters have competences  $\{0.26, 0.26, 1.00\}$ , then the first two will vote to together to override the third more than half the time even though the third is always correct (Vermeule 2009, p. 29, referring to Grofman et al. 1983).

findings in Section III concerning just how far independence can be relaxed without compromising collective group competence too badly.<sup>2</sup>

Of course, there is the further background assumption – rarely stated as such in formal descriptions of the CJT apparatus but clearly at work there – that there is some correct answer to be discovered. Many analysts refuse to apply the CJT to politics, doubting that that is true in the political realm (Black 1958: 163; Miller 1992: 56). Among democratic theorists, it is often said that the "vote of the people" is itself literally constitutive of "the correct answer" politically, rather than there being any independently correct answer that people are supposed to be tracking through their votes. Among legal theorists, too, there are those who would say that there is "no right answer" that judges are supposed to be tracking (Dworkin 1985); the right answer to the question of "what is the law?" is, they would say, just a matter of whatever the judges say it is. For purposes of this paper, however, we shall simply join Sunstein (2009: 10) and Vermeule (2009: 31-3, 67) in presuming that there is some correct answer to the question of what principle of law should apply and what is the legally correct judicial decision in any particular case. That may not always be the case. But it is at least sometimes the case, at least in some respect and to some extent.

<sup>2</sup> There are further worries, which we do not address directly here, that it might not be conceptually possible for the competence and independence assumptions both to be justified (Dietrich 2008).

## II. Traditionalism: A Model of Sequential Decisions

The original CJT set-up tends to assume that voters are making decisions simultaneously with one another, or anyway in ignorance of (or anyway, indifference to) what other voters have done. ~~The Traditionalist~~ approach to constitutional interpretation envisages something very different. There, courts judges are making their decisions sequentially, ~~not~~ simultaneously. Furthermore, when being Traditionalists, subsequent judges make their decisions not only in knowledge of but also in deference to earlier judges' decisions. That changes things dramatically, from a CJT perspective.

Suppose your court is the first court to take the Traditionalist approach to constitutional interpretation. Suppose that before you, there have been 999 judges who have independently judged the same matter. Then even if each of those judges had been only  $p=0.55$  likely to be correct, the majority among the 999 of them would (by standard CJT calculations) be  $Pr=0.999$  likely to be correct. The "many minds argument" would have real epistemic bite, in those circumstances.

---

But why on earth suppose that your generation is the first to implement Burkean principles? Burke has been a famous figure over two centuries, and principles of stare decisis go back long before that.

Consider, then, the opposite extreme. Suppose that only judges on the very first court to consider the question exercised any independent judgment.

Suppose that thereafter all subsequent judges that have considered that same question have simply deferred to that same initial court's decision (or

deferred to previous judges who were deferring to that initial court in turn).<sup>3</sup>

In that case, the probability of each subsequent court reaching the correct solution is exactly the same as the probability that the initial court reached the correct solution (which would be  $Pr=0.621$  assuming each of the 9 judges on the court is individually  $p=0.55$  competent).

Where each member of every subsequent court votes exactly the same way as the initial court, precisely because that was the way the initial court voted, there is no independent judgment being exercised by any members of the subsequent courts. What we would then have is a case, not of "many minds," but rather of "few minds, many mimics." On that scenario, Traditionalist judges would do no better, epistemically, following many previous courts' precedents than they would deciding the matter for themselves. Their probability of being collectively correct is  $Pr=0.621$ , either way on the assumptions just sketched.

In this section of the paper, we will model scenarios in between those two extreme cases of complete independence and complete lack of independence.

---

---

### A. Modeling Sequential Court Decisions

<sup>3</sup> What constitutes the "same question" is of course an important matter of interpretation in and of itself, allowing scope for independent judgment. To simplify modeling, we ignore that.

In our model of Traditionalism, we will assume that judges have to decide on the same dichotomous question at different points in time. We assume that each judge has an independent and symmetric private signal as to which of the two alternatives is the correct one; and we assume that that signal is more likely to point to the correct alternative than the incorrect alternative (analogously to the CJT's competence assumption). We assume that each judge is attempting purely to make a correct decision in the current case. We further assume that, in doing so, each judge forms her decision on the basis of (a) her own private signal and (b) the history of votes in previous courts (which we assume to be common knowledge).

It is well known that any set-up of this kind can easily give rise to informational cascades.<sup>5</sup> That is to say, the history of votes can be such that it provides so strong evidence in favor of one alternative that all subsequent judges will always follow the evidence derived from that history, and never vote according to their own private signals. If that happens, an informational cascade will have begun. Thereafter, judges will have stopped learning from

<sup>4</sup> Assuming average competence within their court is the same as that within the initial court.

<sup>5</sup> The literature on informational cascades has grown rapidly and has by now resulted in all sorts of technical refinements that we will not address in this paper. Seminal papers are: Bikhchandani, Hirshleifer and Welch (1992, henceforth: BHW), Banerjee (1992) and Smith and Sorensen (2000), a useful introduction is Chamley (2004). A very simple setup, roughly in line with the model introduced by BHW, suffices to clarify the problem with Many Minds Arguments based on sequential judgments. In our treatment, we will not always model judges as fully Bayesian rational, unlike BHW and others. In particular, we want to maintain the possibility that judges can be irrationally overconfident about their own private signal, or that they vote only according to their

private signal as a matter of principle. This assumption of “bounded rationality” makes room for the assumptions that may ultimately be more realistic than full Bayesian rationality for the phenomena at hand. Of course this could be done in a Bayesian framework by changing the utility functions of the judges: they might prefer a process in which they vote only according to their private signal, so that they do not (only) care about being right. But the Bayesian

their own signals; instead they will follow blindly the judgment suggested by the voting record of previous courts. An informational cascade is problematic in that, once it begins, the informational base on which all subsequent judgments are based can be very thin, and all information from private signals is thereafter systematically ignored. This is what Vermeule (2009: 75-7) calls the "Burkean paradox."

One natural way to model sequential decision-making would be for a judge to take all historic votes, add her own private signal as yet another vote, and then vote for whichever alternative has a majority among those pooled votes.<sup>6</sup> (After all, if all judges are assumed to be equally competent, it would seem only natural for a judge to count each of every other judge's opinion equal to her own.) But some judges might be more stubborn than that, weighing their own views more heavily than others'. And some judges might have shorter historical memories, or different standards of "how far back it makes sense to go" in looking for precedents.

---

Our model will therefore introduce two further parameters (which we will assume to be identical for all judges, until Section I.C. where that assumption is relaxed). The first additional parameter reflects the weight,  $w$ , that each judge gives to his own private signal relative to the votes of other previous judges. Operationally,  $w$  is equal to the number of "votes" that the

---

treatment comes with some algebraic costs and little gain for the purposes of this paper.  
6 This is the standard set-up in the "informational cascade" literature: see BHW, for example.

judge allocates to himself in reckoning which alternative wins after pooling all previous votes with his own based on his private signal.<sup>7</sup>

The other parameter reflect the length of each judge's memory,  $m$ .

Operationally,  $m$  equals the number of previous decisions they take into account in reaching their own decision. If  $w$  is large (larger than  $m$  times the number of judges on a court) or  $m=0$  the judge will vote on the basis of her own private signal alone. If  $w=1$  and  $m=2$ , then the judge will vote against her own private signal if and only if both previous judges have voted against it.<sup>8</sup> If  $w=2$  and  $m=3$ , then the judge will vote against her own private signal if and only if all three previous judges have voted against it.

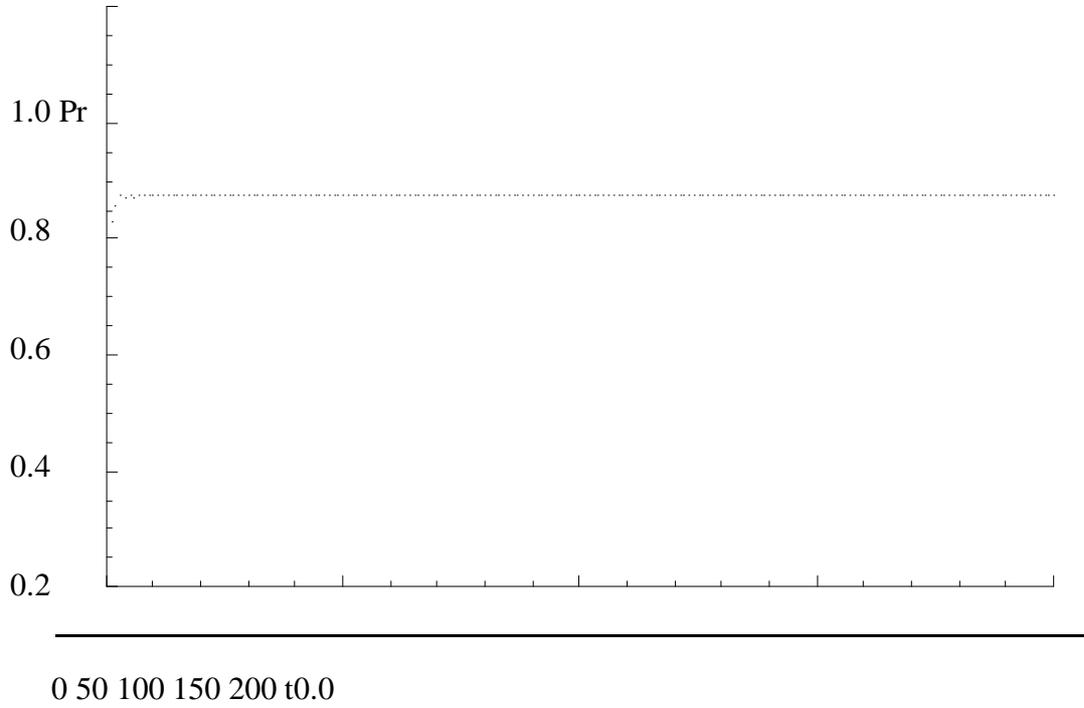
Suppose the competence of each judge (i.e., the reliability of her private signal) is  $p=0.55$ . Suppose that each judge weighs her own private signal equally with every other previous judge's vote ( $w=1$ ) and suppose that each judge's votes taking into account the last two hundred votes ( $m=200$ ). The results of a simulation employing those settings are shown in Figure 1.

[Figure 1 about here]

---

<sup>7</sup> This is one way of formalizing Vermeule's (2009: 76) suggestion that "individual judges might adopt an intermediate approach, according to which they give some but not complete deference to the views of the past, and correlatively think for themselves to some degree or in some circumstances."

<sup>8</sup> We assume that in the case of a tie the judge votes on the basis of her own private signal.



**Figure 1:** Probability of correct majority decision from a court with 9 members, each  $p=0.55$  likely to be individually correct, and weight of  $w=1$  and memory of length  $m=200$

What we see in Figure 1 is a classic cascade. The probability that the current court's decision is correct plateaus at well under  $Pr=0.7$ ; and it reaches that plateau after only a handful of previous decisions. After at most half a dozen previous decisions, subsequent judges are merely playing "follow-the-leader" rather than revealing any independent judgment (private signal) of their own through the process.<sup>9</sup>

---

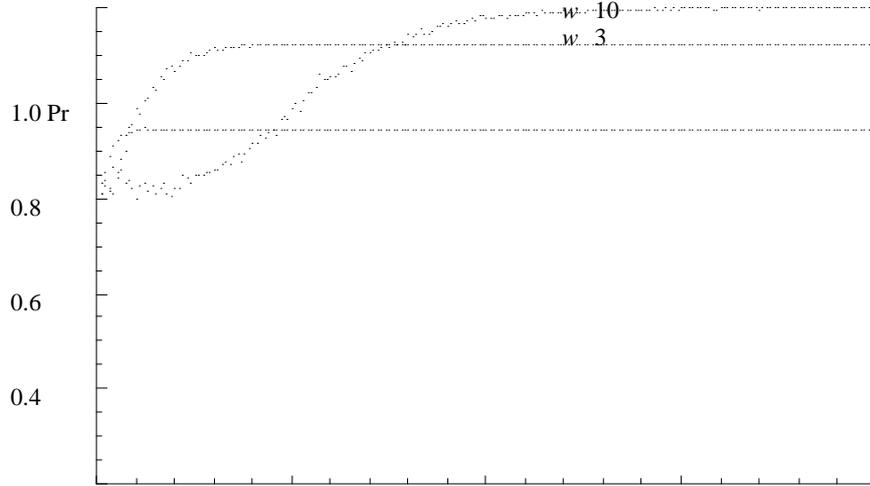
### **B. Solution 1: Stubborn Judges**

Next let us consider the case of "stubborn" judges. Increasing the weight that each judge gives to his own private signal relative to previous judge's votes has two effects. First, it delays the onset of a cascade. Second, in the long term it increases the probability of the current court's decision being correct (although in the short term it might actually reduce that probability, compared to judges with  $w=1$ ). Both of these tendencies are shown clearly in Figure 2, which displays the results of simulations for  $w=3$ ,  $w=10$  and  $w=50$ .

[Figure 2 about here]

---

9 In most cases a single round suffices to start the cascade. Only very rarely will the result be so tight over several rounds that a cascade is delayed.



0.2

0 50 100 150 200 t 0.0

**Figure 2 :** Probability of correct majority decision from a court with 9 members, each  $p=0.55$  likely to be individually correct, and memory of length  $m=200$

These findings admit of easy explanation. Insofar as the previous court has itself taken yet earlier courts' decisions into account in making its own, the amount of "new information" obtained from that court (over and above the information already provided by the decision of the courts that court is following, which we have already taken into account) is strictly limited. After a certain point, reached very soon by judges who treat others' judgments as even roughly on a par with their own (e.g.,  $w=3$ ), an informational cascade of the same sort that Sunstein associated with Populism sets in with the Traditionalist approach as well. Then instead of "many minds" we merely have "many mimics" who confer little epistemic advantage.

The upshot of this analysis is clear. The only way in which courts composed of homogeneous judges will be able to achieve any substantial epistemic success is by judges attaching very little importance to the judgments of previous judges, relative to their own. That is to say, judges would have stubbornly to stick with their own views in the face of a very substantial body of Traditional evidence in the opposite direction. In short, in this scenario, achieving the epistemic power of the "many minds" would require judges to resist Tradition rather than bowing to it.

## **C. Solution 2: Discerning Traditionalists**

In the previous set-up, we assume that judges can only ascertain how previous judges voted, not why. Specifically, that set-up assumes that they cannot determine whether a previous judge has voted predominantly on the basis of her private signal or on the basis of the history of votes of prior judges.

That is not a very realistic assumption. After all, judges can and almost invariably do read the opinions of previous judges. From that, they can usually (albeit imperfectly) surmise the extent to which any given previous judge was following her own judgment or that of prior courts. Let us now vary our set-up to reflect that fact.

For this set-up, we first hypothesize a heterogenous court.<sup>10</sup> Some judges on that court always reveal their private signals. Other judges on that court (indeed, we hypothesize, a majority of judges on the court) make their votes as in the previous set-up, by pooling their private signal with the votes of judges on previous panels. But – and this is the second crucial difference between this set-up and the last – we assume that in so pooling current judges take account only of the votes of judges on previous panels who voted on the basis of their own private signals. That is to say, these judges vote on the

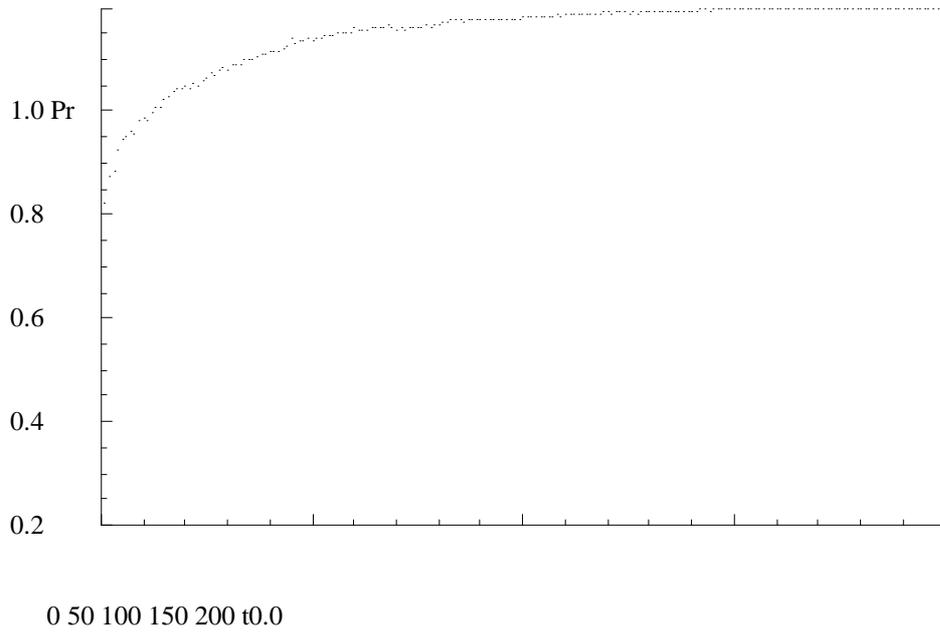
<sup>10</sup> Courts cannot be composed entirely of "discriminating Traditionalists" of the sort we describe below, because "discriminating Traditionalists" require the existence of independent judges stubbornly exercising their own judgment to discriminate in favor of. But on the present account, non-conformism on the bench is of value just because there are others on the bench ("discriminating Traditionalists") who recognize the non-conformists and ignore the opinions of the rest.

<sup>11</sup> Vermeule (2009: 76) anticipates this part of our model, but not the next, when writing, "Perhaps some judges in the stream of precedent or tradition have contributed independently, while some have not..."

basis of previous judges' votes only when those votes are truly "informative" and not merely the product of an informational cascade.

To explore the implications of this sort of set-up, we conduct another simulation. For that, we assume courts consisting of 4 judges who always vote exclusively on the basis of their own private signal and 5 judges who give weight  $w=1$  to their own private signal when pooling it with all such previous informative votes. Figure 3 shows the epistemic performance of courts of this sort, with varying numbers of previous courts of the same sort. As we see there, the probability that such a court will reach the correct decision does not plateau in this case. Instead, it continues to increase the more previous courts there have been.

[Figure 3 about here]



**Figure 3:** Probability of correct majority decision from heterogeneous 9member (each  $p=0.55$  likely to be individually correct, and memory of length  $m=200$ ), where 4 members vote their private signal and the rest decide by weighting their own private signal  $w=1$  and only take into account informative votes.

That probability approaches 1 more slowly, to be sure. As Figure 3 shows, after taking into account 50 previous courts the probability of the majority of the current court reaching the correct decision is only around  $Pr \sim 0.90$ , and after 100 is still only around  $Pr \sim 0.95$ . Still, by the time 200 previous courts have been taken into account, a correct decision is virtually certain. And with courts that are heterogeneous in this way, the probability of a correct decision is a much more rapidly increasing function of the number of previous courts than it is with homogenous courts with very stubborn judges: the curve representing very stubborn judges ( $w=50$  in Figure 2) does not catch up with that representing discerning Traditionalists (in Figure 3) until there have been almost a hundred previous courts. In addition, the heterogeneous courts converge to 1, while courts of homogeneously stubborn judges plateau at a level close to 1 but do not converge to 1.

This, then, is a second way the Traditionalist argument might work. Judges on heterogeneous courts can improve their chances of being correct by taking into account the decisions of previous judges, provided they do so in this very particular way. But note well the irony. Traditionalist courts of this sort benefit epistemically only from judges taking account of the votes of previous judges who were not themselves Traditionalists, and who voted

purely on the basis their own private signal rather than on the basis of the history of votes before them.

### **E. Traditionalism and Cosmopolitanism Share the Same Problem**

So far we have explored the issues of information cascades purely in relation to Traditionalist courts. Notice that just the same problems arise with Cosmopolitan courts as well, however. The Cosmopolitan court is supposed to take account of previous decisions in foreign jurisdictions. But just as there is no reason to suppose that your own court is the very first one to take a Traditionalist stance toward its interpretive task, so too is there no reason to suppose that your own court is the very first one to take a Cosmopolitan stance toward its interpretive task. Yet if the foreign courts from which Cosmopolitan judges borrow have themselves simply borrowed from other, earlier foreign courts, then once again we could easily have a case not of 'many minds' but merely of 'many mimics'. In extremis, all the foreign courts from which you are borrowing might themselves have borrowed (either directly or at one or more remove) from one and the same Ur-court that set the very first precedent that then got picked up in all subsequent decisions across all the different jurisdictions.

There are of course various other problems in implementing a Cosmopolitan approach to constitutional interpretation. One among them,

obviously, is determining which foreign jurisdictions are good comparators to your own. (There is an analogous problem, perhaps, for a Traditionalist in determining which past precedents are good comparators to the case before the current court [Vermeule 2009: 71-2].) No doubt those are important

problems, too. All we want to say, here, is that both Traditionalism and Cosmopolitanism are in any case seriously compromised by risks of informational cascades that are endemic to both of those fundamentally sequential decision processes.

---

### III. Populism: Modeling the Impact of Opinion Leadership

The third 'many minds' approach that Sunstein considers is Populism, in which courts defer to popular opinion.

On its face, that looks like it should be a good plan from an epistemic point of view.<sup>12</sup> The lesson of the CJT is that 'more informants are better' – provided they are more competent than not (better than random; more likely to reach the right conclusion than any wrong conclusion) and provided they are independent of one another.

Worries arise insofar as either of those provisos is not met. If voters are ~~individually more likely~~ to be wrong than right, then the CJT goes into reverse. In the standard binary set-up, if individual voter competence is  $p < 0.5$  then a majority among them is less likely to be right than any one of

<sup>12</sup> Of course, mass electorates rarely vote directly on issues of constitutional interpretation in exactly

the form in which they come before the courts. So judges will have to employ indirect and imperfect mechanisms for surmising where the majority of voters would stand on the matter before the court (Vermeule 2009: 50-3). Still, provided the conditions of the CJT are met, the verdict of a majority of the electorate would be almost certainly right; and it would be epistemically better for judges to follow the electorate than to follow their own best judgment, just as long as judges

them; and the probability that the position supported by the majority among them is incorrect is a rapidly increasing function of the number of voters, approaching certainty as the number of voters approaches infinity.

It may seem hard to imagine how any large number of people can, on average, be worse than random. You would almost have to know the right answer, and then willfully vote against it, in order to do worse than choosing your response at random. A large number of people can be worse than random only if some influence systematically distorts their judgment. One way that might happen would be through heuristics that systematically distort everyone's response in the same way. Another probably politically more important way in which a majority of people might be worse than random is if some outside influence systematically distorts the judgment of many people at once and in the same ways.

In Sunstein's (2009: 125, 169, 171-3, 175) view, the latter is the greatest epistemic worry with the Populist model. Voters might not be independent of one another in the way the CJT requires. Instead, their votes might be subject to the influence of "opinion leaders" who influence many voters at one and the same time and in the same incorrect direction.<sup>13</sup>

~~Formally, the presence of opinion leaders deprives the CJT of a~~  
necessary condition required for its proof to go through. What then happens

are sufficiently better at surmising what the majority of the electorate thinks than they are at surmising directly for themselves the correct outcome.

<sup>13</sup> There are other ways to induce dependence, but we focus on opinion leaders in this paper.

in practice depends on precise parameter settings. That is what we will be exploring through the modeling exercises reported below.

---

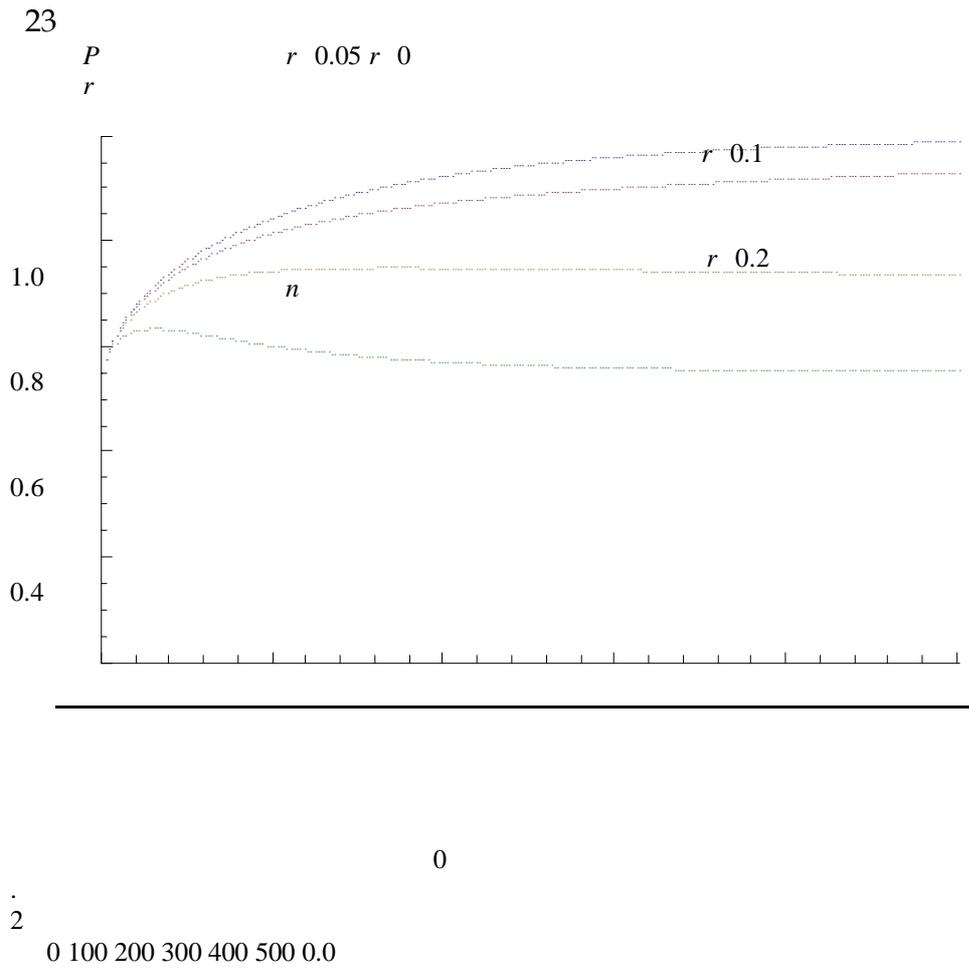
### A. Single Opinion Leader

Let us first demonstrate the epistemic consequences of having a single opinion leader. When we say "R is an opinion leader with respect to voter P" we mean that, once R has announced his position, then P will, with probability  $r$ , adopt the same position as R without considering her (P's) own private signal.

For illustrative purposes, consider the case of a set of voters, each of whom is individually  $p=0.55$  competent (when not following the opinion leader), and each of whom is attached to an opinion leader who is herself also  $p=0.55$  competent. Figure 4 displays curves representing the probability that a majority among the voters will be correct, for various values of  $r$  (the probability of each voter adopting the position of the opinion leader without considering his own signal) and various values of  $n$  (numbers of voters).

---

<sup>14</sup> See Boland et al. (1989) for a technical treatment of this setting.



**Figure 4:** Probability of correct majority decision among voters with individual competence  $p=0.55$ , given a single opinion leader followed with probability  $r$ .

[Figure 4 about here]

As we see from Figure 4, a majority among 180 or more such voters will be  $Pr > 0.9$  likely to be correct if they do not follow an opinion leader at all ( $r=0$ ). If they follow an opinion leader five percent of the time ( $r=0.05$ ), they can still achieve  $Pr > 0.9$  likelihood of a majority among them being correct; it will merely take something like double the number of voters to achieve that. But if they follow the opinion leader ten percent of the time ( $r=0.10$ ), the probability that the majority among them is correct peaks below  $Pr=0.75$  with around 180 voters, and actually begins declining with more voters than that. With voters who are even more likely than that to follow the opinion leader, things get even worse. At best, the probability that a majority among them is correct is only a little higher than the probability the opinion leader is. And as the number of such followers increases, the probability that a majority among them is correct asymptotically approaches  $Pr=0.55$ , corresponding to the probability  $p=0.55$  that the opinion leader herself is correct.

The last-described result suggests why we should rightly be worried about the effects of an opinion leader on the epistemic performance of electorates. If voters follow the opinion leader at all closely, the epistemic power of "many minds" contracts quickly to the epistemic power of the single mind of the opinion leader.

of-following-

Specifically, for  $r < (p-0.5)/p$  - the basic CJT effect still occurs, only a little

o  
r  
e

g  
e  
n  
e  
r  
a  
l  
l  
y  
,  
w  
i  
t  
h

s  
u  
i  
t  
a  
b  
l  
y

l  
o  
w

p  
r  
o  
b  
a  
b  
i  
l  
i  
t  
i  
e  
s  
-

more slowly. Even if everyone follows even the same opinion leader just a little bit, a majority of them will be more likely to be correct than any one of them would be, and that probability asymptotically approaches 1.0 as the number of voters approaches infinity, for sufficiently small values of  $r$ .<sup>15</sup> We see that in the case of  $r=0.05$  in Figure 4, for example. Where the probability of voters following an opinion leader is higher than that – specifically, for  $r > (p-0.5)/p$  – the probability that the majority among them is correct asymptotically approaches the probability that the opinion leader herself is correct. We see that in the case of  $r=0.2$  in Figure 4, for example.

### **B. Two Diametrically Opposed Opinion Leaders**

There may be occasional cases in which there is but a single opinion leader. Really successful totalitarian regimes might be one such case; Berlusconi's Italy might be another. But in most places there are multiple opinion leaders. What difference might this make to the epistemic competence of electoral majorities?

As a first stab at answering this question, let us consider another case – all too realistic – of a highly polarized polity. Suppose for the sake of this stylized example that there are exactly two opinion leaders; and suppose that

the electorate is equally divided with respect to which it follows, with every voter following one or the other of the two opinion leaders with the identical

<sup>15</sup> See the more technical version of this paper for a derivation of this result.

probability. Specifically, for the sake of our example, suppose there are 1001 voters, 501 of which follow Leader A and 500 of which follow Leader B, each with identical probability  $r$ .

The first line of Table 1 reports the probability of the majority among such an electorate reaching a correct conclusion, for varying values of  $r$ . As we see there, that probability is still very high even for moderately high values of  $r$ . For example, even if people are  $r=0.30$  likely to follow their respective leader, the probability that the majority vote is correct remains  $Pr=0.990$ . Indeed, even if people are  $r=0.75$  likely to follow their respective leader, the probability that the majority vote is correct remains  $Pr=0.885$ .

[Table 1 about here]

27	$r$ (probability of followers adopting position of opinion leader)						
	$r=0.1$	$r=0.2$	$r=0.3$	$r=0.5$	$r=0.75$		
group partition							
501, 500	0.997	0.995	0.990	0.966	0.885		
	0.974	0.774	0.572	0.550	0.550		

701, 301

Table 1: Probability that a majority of voters would be correct if they followed two diametrically opposed opinion leaders, to varying degrees.

It is easy enough to see why that should be so. Take the case of  $r=0.1$ . There, we expect about 50 voters follow Leader A and 50 voters to follow diametrically opposed Leader B, and those expected votes would be expected roughly to cancel one another out. The result of the election, then, is basically determined by the votes of the remaining 901 or so voters who do not follow the lead of either opinion leader, and exercise truly independent judgment of the sort the CJT relies upon. The consequence of opinion leadership, in that scenario, is in effect to reduce the size of the useful electorate – from 1001 to something more like 901 in that case. But of course as long as the remaining number of genuinely independent electors is large enough, no matter.

Notice, however, that that result crucially depends upon a presumption of symmetry in the size of each leader's potential followership. In the example given in the top row of Table 1, we assume that there are (almost) equal numbers of voters potentially following each opinion leader, and that each follows his or her respective potential leader with the same probability. Any asymmetries in either respect can upset that result dramatically, as the second line of Table 1 demonstrates. There we report what would happen if substantially more voters are likely to follow one leader rather than another: no (anyway, not much) matter, if the probability

of their following her is low; but if they are likely to follow her with higher probabilities, the epistemic competence of majority voting is seriously compromised where the leaders' respective potential followerships are of appreciably different sizes. A similar result would obviously obtain if the

probabilities-of-following differed substantially between different leaders, as well.

---

### **C. Multiple Uncorrelated Opinion Leaders**

The argument of the previous section is that the electoral effects of diametrically opposed opinion leaders would more-or-less cancel one another out (or, more precisely, the votes of those who follow them would). Consider next the case of opinion leaders whose positions are completely uncorrelated with one another. What would happen then?

Two effects play a role. First, obviously, the more opinion leaders we have, the more independent and competent points of judgment that determine the majority vote, increasing the group competence. But second, and more to the present point, the votes of voters influenced by the opinion leaders might once again roughly cancel each other out.

To see this, suppose there are a great many opinion leaders (as many as there are voters would be the limiting case), and these opinion leaders are wholly independent of one another. For the limiting case, suppose the opinion leaders were incompetent, i.e. not better than random ( $p=0.5$ ) and

distribute themselves utterly randomly between the two propositions at stake

- so they are wholly independent of one another, and there is no correlation

between the positions they took. Further suppose that each opinion leader has the same number of voters following him with the same probability.

Then the expected votes caused by opinion leaders would roughly (if rarely exactly) cancel each other out, and the remaining independent votes would by and large dictate the outcome of the election. And of course that is the worst-case scenario. If opinion leaders are more competent than that, then they tend to produce more followers for the correct alternative, pushing those who follow them in the right direction more often than not as well.

In the real world, of course, we do not have as many opinion leaders as voters. So the real-world question is just how many uncorrelated opinion leaders would suffice for those effects to be more-or-less attained. That is a computational question, to which Table 2 offers some sample calculations as an answer.

[Table 2 about here]

$r$ (probability of followers adopting position of opinion leader)						
Opinion leaders	Group partition	0.1	0.2	0.3	0.5	0.75
1	1001	0.72	0.55	0.55	0.55	0.55
3	333, 334, 334	0.87	0.80	0.63	0.58	0.57
5	1x201, 4x200	0.93	0.85	0.77	0.60	0.59
11	11x91	0.99	0.93	0.85	0.74	0.64

**Table 2:** Probability of correct majority decision among 1001 voters with individual competence  $p=0.55$  split evenly among multiple uncorrelated opinion leaders (also  $p=0.55$ ), following his respective opinion leader with probability  $r$ .

From Table 2 we can see the effects of even a modest number of uncorrelated opinion leaders with same-sized potential followings. Of course everything depends heavily, here as before, on just how likely people are to follow their respective opinion leaders. But even if they follow their respective opinion leader with a probability of  $r=0.5$ , where there are eleven wholly independent opinion leaders whose positions are utterly uncorrelated with one another's, the probability that the majority of the electorate is correct will still be (on the scenario envisaged in Table 2) be almost 75%.

---

#### **IV. Conclusion**

The upshot of these computational exercises is that, from an epistemic point of view, Traditionalism is worse and Populism better than they initially appeared to Sunstein. Traditional courts can succeed epistemically only under the most ironic of conditions: either if the judges on them do not pay very much attention at all to previous decisions, or if judges on them pay attention only to previous judges who did not pay much attention to previous decisions.

Populism can fail epistemically if there are only a few very strong opinion leaders; but if there are several of them with roughly equal-sized potential

followerships then that does not compromise the epistemic power of popular majorities very badly. The "many minds" argument, properly applied, would

be more in favor of Populism and less in favor of Traditionalism than Sunstein supposes.

**References**

- Banerjee, A. V. 1992. A simple model of herd behavior. *Quarterly Journal of Economics* **107** (3): 797–817.
- Bikhchandani, Suhil; David Hirshleifer; and Ivo Welch. 1992. A theory of fads, fashion, custom and cultural change as informational cascades. *Journal of Political Economy* , 100 (5): 992-1026.
- Black, Duncan. 1958. *The Theory of Committees and Elections* . Cambridge: Cambridge University Press.
- Boland, P. J.; F. Proschan; and Y. Tong. 1989. Modelling dependence in simple and indirect majority systems. *Journal of Applied Probability* **26**(1): 81-8.
- Chamley, C. P. 2004. *Rational Herds: Economic Models of Social Learning*. Cambridge: Cambridge University Press.
- Dietrich, F. 2008. The premises of Condorcet's jury theorem are not simultaneously justified. ~~*Episteme* **58** (1): 56-73.~~
- Dworkin, Ronald M. 1985. Is there really no right answer in hard cases? Ch. 5 of Dworkin, *A Matter of Principle*. Cambridge, MA: Harvard University Press.
- Estlund, David. 2008. *Democratic Authority: A Philosophical Framework* .

Princeton, NJ: Princeton University Press.

Grofman, B.; G. Owen; and S. L. Feld. 1983. Thirteen theorems in search of the truth. *Theory & Decision* , 15: 261-78.

- Owen, G.; B. Grofman and S. L. Feld. 1989. Proving a distribution-free generalization of the Condorcet Jury Theorem. *Mathematical Social Sciences*, 17: 1-16
- List, Christian and Robert E. Goodin. 2001. Epistemic democracy: generalizing the Condorcet Jury Theorem. *Journal of Political Philosophy*, 9: 276-306.
- Miller, David. 1992. Deliberative democracy and social choice. *Political Studies*, 40 (#5: Special Issue): 54-67.
- Smith, L. and P. Sorensen . 2000. Pathological outcomes of observational learning. *Econometrica* 68(2), 371-98.
- Sunstein, Cass R. 2009. *A Constitution of Many Minds: Why the Founding Document Doesn't Mean What It Meant Before*. Princeton, NJ: Princeton University Press.
- Vermeule, Adrian. 2009. *Law and the Limits of Reason* . Oxford: Oxford University Press.