

University of Chicago Law School

Chicago Unbound

Public Law and Legal Theory Working Papers

Working Papers

2022

Appellate Panels of One

Saul Levmore

Follow this and additional works at: https://chicagounbound.uchicago.edu/public_law_and_legal_theory



Part of the [Law Commons](#)

Chicago Unbound includes both works in progress and final versions of articles. Please be aware that a more recent version of this article may be available on Chicago Unbound, SSRN or elsewhere.

Appellate Panels of One

Saul Levmore

Introduction

Appellate review can be understood as an opportunity to correct errors made by lower courts and, by virtue of their multi-member panels, as a way to benefit from the wisdom of crowds. Appellate panels of three judges, and then a larger Supreme Court of nine, are likely to interpret, apply, or advance law more correctly, or simply better, than a single lower-court judge whose effort is under review.¹ Appellate review can also be understood as relying on more experienced or otherwise superior decision-makers or as designed to make law more uniform, inasmuch as lower court decisions on many matters will converge as they follow precedents. It has also been understood as a way to take advantage of the knowledge that litigants themselves have about lower court errors.² Appellate review is surely a means of encouraging more careful work by lower courts; people are often more careful when they know that their work can be reviewed or observed by superiors or well-regarded peers. Finally, appellate review may be of great value (even) when it affirms a lower court, because each step adds to the development of a lasting precedent. Most of these perspectives have counterparts in other settings where the familiar question of when to seek and pay for a second opinion arises. But most second opinions, whether sought before agreeing to a medical procedure or contracting for an auto repair, are given by a single analyst, while appellate review in the federal and most state systems normally involves three jurists, and then yet more in the event of a further appeal.³ This Essay examines the logic of second and third opinions – even without the added complexity introduced by the precise cost of review (in the form of time or money) – and reaches

* I am grateful for the discussions of probability theory I enjoyed with Dylan Baker and the rewarding exchanges I had with Frank Easterbrook, Dylan Fane, and Eliot Levmore. Comments welcome at slevmore@uchicago.edu

¹ Error correction, uniformity of law, and a redirection of the role of the Supreme Court are thought to explain the Evarts Act, and the creation of the U.S. Federal Courts of Appeal in 1891, not to mention the growth – in number, size, and caseloads – of these appellate courts. See Chapter 1, *A Brief History of the Circuit Courts* in William M. Richman & William L. Reynolds, *INJUSTICE ON APPEAL: THE UNITED STATES COURTS OF APPEALS IN CRISIS* (2012). Error correction – whether in federal or state courts – is the goal pursued here, though uniformity and other aims can be understood as compatible with correct decisions. The meaning of “correct” is discussed in Part IV. When decisions are viewed as correct, precedential value and uniformity become especially attractive.

² Steven Shavell, *The Appeals Process as a Means of Error Correction*, 24 *J. LEGAL STUD.* 379 (1995).

³ When private parties opt for arbitration, they are not aiming for something equivalent to appellate review. They usually deploy one or three arbitrators, and none of the players is establishing judicial precedents. In the conventional legal system, the costs of appellate review is partly borne by the government and might overused (or overused), but the optimal rate of review is not the subject here.

several counterintuitive results. Most appellate processes should be restructured so that one judge alone reviews the lower court. Only if this single appellate judge disagrees with the lower court, should one more judge enter the fray, and even that may be wasteful. Legal questions that are appealed will normally be decided by 2-0 or 2-1 decisions, involving just one or two appellate judges in addition to the lower court judge.

There are other reasonable conclusions to reach once the logic of appellate review, and second opinions quite generally, is examined. There is a case to be made for having the first appellate judge always decide whether further review is in order. On the other hand, and to the contrary, the appellate process could always stop after one review – subject to the Supreme Court’s deciding to take the case. These and other possibilities are examined here, but mostly set aside in favor of the central argument about the appellate process. The arguments that drive the conclusions are fueled by some probability theory, and have surprising implications for areas outside of law in which second opinions are commonly sought. I begin with the idea that reaching the *correct* decision (defined presently) is the immediate goal. As the argument proceeds, the value of long-lasting rules and other aims are brought into play.

Part I begins with the most familiar use of second opinions. It rethinks the wisdom of soliciting another assessment before following a recommendation regarding a serious medical intervention. There are important differences between medical and legal decisions, but it is instructive to begin with an example where it is easier to insist that there is a correct answer. The analysis shows that the common thinking about the value of a second medical opinion is poorly conceived. The goal in Part I is not to give medical advice, though there is some of that, but to rethink the logic of second opinions, before applying it to judicial review. Part II then takes account of some of the ways in which judicial review is unlike other calls for review. It suggests that if we incorporate the likelihood that a lower court judge is correct, it is sensible to move to a system where we begin, and usually end, with a single appellate judge. Part III tests the idea of a single reviewer by looking not only at the value of discussion and teamwork among judges, but also at the importance of some assumptions made here about the probability that a judge is correct. Part IV turns to the possibility that appellate judges are not deployed to find correct answers, perhaps because there is often no such thing, but rather to reflect and aggregate preferences, a distinction familiar to students of public choice, not to mention to observers who regard much of what judges do as reflecting political preferences. Part V extends the analysis of appellate review to committees and boards of directors, and then to juries. The insights offered here suggest some changes in law, though some of these are likely to be politically impossible in the near future.

I. Second Opinions

When a second opinion can be obtained from someone who is much more likely to be correct than the first analyst, it is a good investment to acquire that second opinion unless it is relatively costly

when compared to the importance of the decision to be made. It is, for example, common for an ailing individual to go to a nearby and readily available doctor, but if that professional suspects that there might be a serious problem, he or she will normally recommend a specialist in the area of concern. We can think of this specialist as the first well-informed decision-maker, because in modern times it is unusual for the family doctor or gate-keeper to move ahead with a serious intervention without consultation. Things are different in law because the first, or lower court, judge is often as specialized as an appellate judge. In medicine, if the gate-keeper recommends an average specialist, but more experienced and skilled specialists are available, then it is obvious that opinions by two specialists might wisely be sought. If the first specialist is 70% likely to be correct in recommending for or against a serious and invasive procedure, but a true expert is 90% likely to be correct, then the value of the second opinion is apparent, and especially so if the first recommends something drastic. Indeed, many health insurance companies encourage, or require and pay for, second opinions in specified circumstances.⁴ There is an equilibrium. When a 90% (probability of being correct) second opinion will be available, it is less important to have a highly skilled first opinion. From a systemic perspective, it can be efficient to train a larger set of simply competent first responders.⁵ Correspondingly, law sometimes pairs a low-cost first decision-maker with a single, and presumably superior, reviewer, who can affirm, modify, or overturn the first. A magistrate's decision is normally reviewed by a single, Article III, judge, much as work by a new associate at a law firm is reviewed by a more experienced attorney. The magistrate comparison is imperfect if only because magistrates, along with hearing officers and arbitrators, are often used as factfinders in settings where it is expensive or impossible for later reviewers to access the same facts. But the larger point is that law already uses first responders who can be reviewed by a single party, thought to be more likely to be correct, when it comes to interpreting and applying law to the facts. Here, as in most of the discussion that follows, it is assumed or understood that (in addition to factfinding) law is looking for a *correct* answer, rather than for the *preferences* or political inclinations of a judge regarding the interpretation and application of statutes, precedents, the common law, or constitutional provisions. These are not the only two things that law might be

⁴ Robert B. Grant, *Outpatient Surgery: Helping to Contain Health Care Costs*, 115 MONTHLY LABOR REV. 33, 34 (1992) (explaining that insurance cost-containment measures include requiring second opinions for specific surgical procedures, and reducing coverage where participants do not seek a second opinion).

⁵ The efficiency claim runs in both directions. The better the appellate court, the more efficient it is to invest less in the lower court. On the other hand, the more we can save costs by having a good lower court, or first doctor, and then a smaller number of reviewers, as argued in this Article, the more sensible it is to invest in good first responders. An additional complication is that while many of us are more careful when we know we might be reviewed, there is the danger of sloppiness at step one because the worker knows that errors are likely to be caught at the next step. This is likely to be a problem when the first worker is not identifiable. In any event, the problem of duplicative review, in the form of suboptimal effort by a given participant, does not appear to be present when judges are identifiable; lower court judges are more likely to fear reversals than to count on appellate courts to correct their errors.

trying to do, and the definition of “correct” is contestable, and addressed in Part IV. Again, most of the discussion here does assume that there is a correct answer to be found.⁶

To simplify the analysis, imagine that specialist #1 recommends that a procedure be undertaken as soon as possible. If the advice is correct, the patient will live or enjoy some other great benefit; if it is incorrect, the medical intervention will lead to disaster. In order to avoid unnecessary arithmetic here, I will imagine that the cost of being wrong matches the benefit of acquiring and following correct advice. For further convenience, assume that the first specialist gives the correct advice 2/3 of the time, and that this probability of a correct diagnosis is the same for any additional specialists who will be consulted. This is a good assumption for the argument presented here, both because it is difficult for the patient (or even a family physician) to know the probabilities to attach to each specialist and because, returning to law, legal questions will normally be reviewed by judges who are randomly assigned to panels, so we can imagine that the average rate of being correct is 2/3 for each additional judge. If there is time and money (or no cost) to get additional opinions before agreeing to an invasive procedure, the well-known “wisdom of crowds,” or Condorcet Jury Theorem, prevails.⁷ The chance of error and a very bad outcome if the patient goes to specialist #1 alone is 1/3. But if the patient goes to three specialists, whom we can refer to as

⁶ When judges are trying to be correct but there is simply no right answer to be found, and it is also not a matter of preferences, it is easy to argue that fewer reviewers (at lower cost) are no worse than more.

⁷ Condorcet thought about jury trials where the defendant was either guilty or not. It is no accident that he wrote at the time of the French Revolution and, indeed, Condorcet himself spent some time in hiding and in prison during those turbulent times. See Cheryl D. Block, *Truth and Probability—Ironies in the Evolution of Social Choice Theory*, 76 WASH. U. L. Q., 975, 999-1000 (1998). Condorcet’s argument is straightforward. If every juror, or voter, is (a) more likely than not to be correct, and (b) these jurors are independent, so that there is no danger of an incorrect voter tarnishing the probability of others being correct, then (c) as the number of voters increases, and (d) a majority voting rule is applied, the probability of the group reaching the correct decision approaches 1. Here and throughout, the assumption is that the decision-makers assess matters independently. The case of three voters is examined at length in note 9, where it is clear that three are better than one. A 3-0 vote is then convenient. In a more extreme case, imagine fifteen voters, be they jurors, judges, or doctors, with each having a probability of .6 of being correct. With a majority rule, we need to find the probability that at least 8 will be right or, put differently, we need to be concerned about any case in which more than 7 are wrong. With a unanimity rule, it is easy to see that the chance that 15 out of 15 are wrong is extremely low (.4 to the 15th power, or vastly less than 1%). It is also the case that a 15-0 correct decision, requiring every single juror to have a good day, is only more likely than not when each juror is more than .95 likely to be right or when a correct, confident, or charismatic juror simply browbeats others or is followed by them. Without compromise (a topic discussed in Part V), most juries of substantial size will be hung under a true unanimity rule. But as for majority rule, and returning to medical opinions, the chance of getting 8 or more incorrect doctors out of 15 doctors, each with a 60-70% chance of being right, is about 21%. This is much better than the error rate assigned here to the single doctor. It is easy to explain Condorcet (and his followers) nearly brushing aside this topic because he was personally interested in unanimous jury decisions with the death penalty in view.

For a clear but also critical view of the Condorcet Jury Theorem as applied to law, see Paul H. Edelman, *On Legal Interpretations of the Condorcet Jury Theorem*, 31 J. LEGAL STUD. 327 (2002) (finding the theorem “unassailable” but critiquing some of its applications to legal questions).

doctors #1, #2, and #3, and sensibly abides by a majority “vote,” as recommended by the Condorcet Jury Theorem, there is an $8/27$ chance that all will agree and be correct ($2/3 \times 2/3 \times 2/3$) and a $1/27$ chance that all three doctors will be wrong ($1/3 \times 1/3 \times 1/3$). A 3-0 decision is thus reassuring and far more likely to be correct than not, but it has not otherwise added anything to the decision-making process. The patient might as well have saved time and money by following #1’s advice. The extra opinions invoked the wisdom of crowds with a 3-0 vote, but in an important sense little or nothing was gained, and of course the added doctors increase health care costs.

Return then to the value of going for that second opinion, rather than comparing three against one, which is to say a unanimous (small) crowd against the individual assessor. If #2’s advice contradicts that obtained from specialist #1, the patient has new information. Again, if #2 agrees with #1, nothing practical is gained from this knowledge. The patient was willing to lose time and money to get added comfort, but there is a negative impact on other people who could be helped by the excess medical advice consumed by the first patient. Indeed, the ability of wealthier patients to obtain second and third opinions, might explain part of the reason that wealth is correlated with longevity. There is a case to be made for the government paying for second and third opinions for less fortunate people, if the private market demonstrates that more privileged people pay for these opinions when on their own.

Even when a patient finds that #2 agrees with #1, the patient might continue on to #3 (we might conveniently assume that at that point money or time will run out) and obtain a 3-0 or 2-1 result but, again, the patient could have stopped after getting #1’s advice. The benefit, in terms of actually deciding whether to submit to the recommended medical intervention, must be in the case where #2 disagrees with #1. If in the face of this 1-1 set of assessments, the patient is unsure how to proceed, and so flips a coin, the patient is no better off than stopping after #1, and might just as well have proceeded without seeking advice.⁸ As a practical matter, #2’s input might be favored, because #2 can look for flaws in #1’s analysis, or build on the earlier work. This is a familiar second-mover’s advantage, even if it is contrary to Condorcet’s technical preference for independent judgments. But the discussion here tries to eliminate this debatable point, first by assuming that #1 and #2 have the identical $2/3$ chance of advising the patient correctly, and then by appealing to many readers’ intuitions, and following the approach taken by Condorcet himself: When each voter, or decision-maker, is more than 50% likely to be right, we must amalgamate *independent* judgments in order to guarantee the theorem’s validity. The idea is not simply to

⁸ This may seem puzzling, inasmuch as the patient has useful information after #1 and then it is updated by the acquisition of more information. How could the acquisition of information make someone worse off? But there is no paradox, because the second piece of information taught the patient that what she thought after hearing from #1 was not as reliable as it first seemed. The second piece could, after all, have confirmed the first. Put differently, the conflict between #1 and #2 tells the patient to seek #3, if possible, and this is informative. It remains true that going to #2 is pointless in terms of outcomes, unless it will be possible to go to #3. Comparing “average accuracies,” as done presently in note 9, is statistically useful but sometimes of no help to the patient.

justify the aggregation of probabilities but, presumably, to avoid the danger of one person's mistake infecting another's judgment. In the case of doctors, as with auto-repair shops, it is common for patients who seek a second opinion to think it is more valuable when #2 is unaware of #1's assessment. On the other hand, and looking ahead to judges, the appellate court is normally thought to benefit from having access to the lower court's reasoning, not to mention the views of fellow panel members. These contrasting intuitions form a subject returned to below.

At this point, the analysis appears to have generated two conclusions. First, when one is advised to seek a second opinion, it is really a suggestion to be ready to seek a third – unless the second opinion is to be acquired from a source significantly more likely to be correct than that which produced the first opinion. If there will be no time to get a third opinion then, counterintuitively, it is usually wasteful to solicit a second opinion. Even the idea that second opinions, in the case of a medical intervention, might provide peace of mind is questionable, inasmuch as when #2 disagrees with #1, the patient is likely to be stressed. Second, and more striking, even when there is time and money for *two* more opinions, one might as well forego the second opinion because after a 1-1 split, the third opinion is no more likely to be right than the first. After a split opinion, the patient is relying on the quality of #3, and by hypothesis this is identical to that already contained in the advice previously received from #1. If the convenient assumption that each decision-maker is equally ($2/3$) likely to be correct is unrealistic, it is sufficient to imagine (as suggested earlier) that the patient has no way of knowing which doctor is superior. Looking ahead to an analysis of legal decisions, an advantage of appellate review might be that an appellate judge is more likely to be correct than the lower court judge, but there is no reason to think that a second or third appellate judge is more likely to be right than a prior appellate judge. This is part of the argument for a single appellate judge rather than a panel of three. To be sure, “correct” is likely to be an appealing goal with respect to medical care, but it might be harder to define where judging is concerned. For now, “correct” is taken to mean that lower courts ought to find the answer that would be found if the appellate process ran its full course. This saves time and money. A fuller discussion of what it means for a judge to be correct is offered in Part IV.

In fact, the second conclusion just stated is mostly wrong, and it is so in a way that reveals an important point about additional opinions, whether by doctors or judges looking for right answers. Counterintuitively, and as already stated, getting a third opinion when one faces a 1-1 split is no more likely to increase the chance of correctness than that available from a 1-0 decision, because it depends on the decision of the tie-breaker, who is no more likely to be correct than was the first assessor. It is the search for a second and then possibly a third opinion that makes the venture worthwhile. The advantage of a second opinion is that if it conflicts with the first, there is the opportunity to obtain an assessment from #3, and then if #2 and #3 agree, they *are* more likely to be right than was #1 alone. Put differently and with the numbers offered here, one opinion has a 67% chance ($2/3$) of being right, but three people, each with a $2/3$ chance of being right, and

governed by majority rule, have about a 74% chance of being right.⁹ Essentially, the reason to go for a second opinion is that it gives one the chance to go for a third opinion and, in this way, get the wisdom-of-the-crowd advantage of three “voters” rather than one or two. A second opinion is of no use on its own. If the first two agree, one might as well have stopped with #1 and, if they conflict, then one is certainly no better off with this tie vote. If there is neither time nor money for more than two opinions, it is sensible to halt after opinion #1.¹⁰

There is no sleight of hand here. Put in terms of the numbers used here, though any reasonable ones will do, it might pay to go to #2, assuming there will be time for #3 if needed, because there

⁹ There are several ways to see this striking point. If the first doctor is correct, the chance of it being overruled by the other two is very low ($1/3 \times 1/3 = 1/9$, if each is $2/3$ likely to be right). And if the first doctor is wrong, the chance of being overruled by the next two is fairly high ($2/3 \times 2/3 = 4/9$). The first doctor is, by assumption, more likely to be right than wrong, but $2/3$ (the chance #1 is right) \times $1/3$ (the chance that #2 is now wrong) \times $1/3$ (the chance that #3 is also wrong, and #2 and #3 have overruled #1’s correct diagnosis) is $2/27$. That is less likely than the chance that #2 and #3 will overrule an incorrect decision, as that chain of events is $1/3$ (#1 is wrong) \times $2/3$ (#2 now gets it write) \times $2/3$ (#3 agrees with #2, and the two of them overrule #1) = $4/27$. This difference (between $4/27$ and $2/27$ here) shows the advantage of going ahead for a second and then sometimes a third opinion.

Another way to see this is to examine all the possible patterns, assuming there is time (and money) for three opinions:

- (a) With $8/27$ probability, there is a correct 3-0 decision ($2/3 \times 2/3 \times 2/3$).
- (b) With $1/27$ chance, they again agree, but are all wrong ($1/3 \times 1/3 \times 1/3$).
- (c) There is a $4/27$ chance that #1 and #2 get it right, and #3 is wrong ($2/3 \times 2/3 \times 1/3$).
- (d) Similarly, there is a $4/27$ chance that #2 and #3 get it right, after #1 was wrong ($1/3 \times 2/3 \times 2/3$).
- (e) There is also a $4/27$ chance that #1 and #3 are correct, and #2 is now wrong ($2/3 \times 1/3 \times 2/3$).
- (f) There is a $2/27$ chance of two incorrect decisions by #1 and #2, while #3 is right ($1/3 \times 1/3 \times 2/3$).
- (g) There is a $2/27$ chance of a bad decision, if #1 and #3 are wrong, and #2 correct ($1/3 \times 2/3 \times 1/3$).
- (h) Finally, there is a $2/27$ chance that #1 is correct and that #2 and #3 ruin things because they are wrong ($2/3 \times 1/3 \times 1/3$).

If the first two doctors agree, the average accuracy is 0.8; this is lines (a) and (c) compared to lines (b) and (f), with 12 correct and 3 incorrect chances (numerators) out of 15. In contrast, if the first two disagree, the average accuracy is .67 – lines (d), (e), (g), and (h). The numerators in those four patterns add to 12, with 8 correct and 4 incorrect. This $8/12$, or .67, is unsurprising because everything depends on #3 who is $2/3$ likely to be right, just like #1. Overall, there are 20 correct results (the numerators in lines (a), (c), (d), and (e)) out of $27 = .74$ or so, which is better than stopping after #1, with his .67 chance of correctness.

Readers with time on their hands might also see that the analysis uses $2/3$ in order to come close to the maximum benefit of three voters rather than one. With different numbers, it is even less advantageous to solicit additional opinions. As the likelihood of correctness drops closer to .5, it is less likely that additional reviewers will improve things, and if the likelihood rises toward .99 then reviewers are again unlikely to improve things much, because the first doctor is so likely to be correct in the first place. Competent or exceptional review is most helpful in the middle.

¹⁰ If the assumptions seem absurd, it is because the reader is imagining a case where the first doctor immediately turns things over to a specialist. In this case we might think of the first doctor as #0, and the first specialist as #1. A second opinion is often recommended after this first specialist recommends a significant intervention.

is a $2/27$ chance of improvement.¹¹ That is not much; it is likely that soliciting a second opinion is normally wasteful. An insurance company that does not pay for the second opinion might defend itself with the preponderance-of-the-evidence rule.¹² Still, a second opinion is worth something, but only if it is understood as opening the door for three opinions rather than one. Doctors #2 and #3 are more likely to correct #1's error than are they likely to incorrectly overrule the good advice provided by #1.

Once it is clear that it is not that a 2-1 decision is more reliable than a 1-0 decision, but rather that three opinions are better than one, the attentive reader will question the value of a 5-4 decision, or for that matter a 50-49 decision. Certainly, a 2-1 decision will be improved if checked by nine more assessors, as in the case of the Supreme Court. If it decides a matter with a 5-4 decision we might be no better off. But overall, the wisdom of nine will be better than that of three, just as three did better than one. There is, however, a catch here. As the number of voters grows, be they doctors or judges, and a close vote is produced, we must reconsider the assumption that each assessor is $2/3$ (or some such probability) likely to be correct. A 5-4 decision is just fine because it is not very far from the assumption of $2/3$ – a 6-3 decision would be perfectly in line with that assumed probability that a single decision-maker is correct. But a 50-49 decision is not of interest simply because it raises the question of whether this is more or less reliable than a 1-0 decision, but rather because it tells us that it is virtually certain that many voters were not $2/3$ likely to be right; it is more likely that many of these voters had no idea what they were doing. A very close result with large numbers is most consistent with the idea that each had about a 50% chance of getting the

¹¹ As we have seen, it will be of benefit when #1 is wrong and then #2 and #3 are correct ($1/3 \times 2/3 \times 2/3 = 4/27$), but it will ruin things if #1 is right and now both #2 and #3 err ($2/3 \times 1/3 \times 1/3 = 2/27$). Alternatively, we can compare .67 to .74 in note 9.

¹² Of course, a well-done defense here needs to take account of the cost of a second opinion, the probability of needing a third opinion (and its cost), then the benefit of a correct decision, and then multiplied by the probability that the review be worthwhile – that it will do more good than harm ($2/27$ in this example). This is unlikely to be standard practice, unless there happens to be a recovering statistician on the jury, on the bench, or even in the insurance company's boardroom. If so, there is room to be concerned about the use of the preponderance of the evidence rule as presently constituted (a topic beyond the scope of this Essay). The Learned Hand Rule may get the negligence calculation just right, for example, but it is doubtful that juries think much about the probability that a given precaution will work, and especially so when it is often useless when combined with another precaution, as is going to #3 when #2 and #1 agree. Even when one's own life is at stake, two concurring opinions seem more than enough for most people, even when they concur for different reasons, and this suggests that something like the preponderance rule is in our minds rather than a probabilistic rule with many values to be estimated. It may also, perhaps oddly, be embedded in the minds of insurance company executives. There does not seem to be an example of an insurance company that suggests, requires, or pays for a third or fourth opinion after #1 and #2 agree, though given the high value of a human life, or the cost of paying for lifelong health care in the event of a disaster, it is easy to insert numbers that make additional opinions worthwhile after a second opinion agrees with the first. Additional opinions can also confuse things when multiple opinions conflict.

matter right.¹³ We can be very confident, when ten doctors who review a case unanimously agree, 10-0, on what the patient should do next. A 7-3 vote might reassure us that our 2/3 estimate was close to the mark, but the patient is less interested in admiring the chosen prior than in getting the correct advice.

The discussion in this Part has demonstrated that there is some wisdom of crowds even if a 2-1 decision might simply mean that the third voter decided things, and was no more likely to be right than the first voter (or doctor). In the case of law, it is noteworthy that an even number of jurors, like 12 or 6, is sometimes (in civil trials in many states) asked for a simple majority vote, and this produces at least a 2-vote margin.¹⁴ It is easier to see that 7-5 is better than 1-0, than it is to believe that 11 jurors were worth empaneling even when they reach a 6-5 result.

II. Second and Third Opinions in the Judiciary

There are two good reasons not to jump to the conclusion that since three opinions are better than one – even if just modestly so – it is correct to have three judges deciding a case on appeal. The primary reason draws on the role of the lower court judge, and the second takes us to the difficult question of how to think about a majority’s decision that is the product of different reasoning by the several participants.¹⁵

When a lower court opinion is appealed, and a circuit court with three judges affirms with a 3-0 vote, we can be confident that the right result has been reached, because three, or really, four judges

¹³ Another complexity is that a close vote may reflect the fact that when voters know that a simple majority will suffice, the cases they agree to hear (or, in legislatures, the bills that are pushed forward) will slide toward a close vote, reflecting the idea of “minimum winning coalitions,” or what was once called “the size principle.” See William H. Riker, *THE THEORY OF POLITICAL COALITIONS* (1962). The discussion in the text avoids this by focusing on binary decisions like whether or not to take on the risk of a medical procedure. However, the passing reference to a 5-4 decision in the Supreme Court must take case selection into account; voters on the extreme may encourage the Court to take on the strongest cases for which a majority vote can be assembled. The idea of an “extremist” might smack of preferences, but it is not entirely inconsistent with the search for right answers. On the one hand, voters on the extreme should be happy to follow the wisdom of the crowd but, on the other hand, they may genuinely think that there are issues regarding which they are more likely to know the correct answer. The matter is discussed further in Part V.

¹⁴ Once we leave the requirement of unanimity, other issues arise. For example, one majority might vote for liability but a different majority might vote for the particular determination of damages. See Shari Seidman Diamond et al, *Revisiting the Unanimity Requirement: the Behavior of the Non-Unanimous Civil Jury*, 100 NW. U. L. REV. 1, 9 (2005) (suggesting the possibility of excluding jurors who voted against liability from participating in the vote on damages).

¹⁵ See generally, Maxwell L. Stearns, *Should Justices Ever Switch Votes?: Miller v. Albright in Social Choice Perspective*, 7 SUP. CT. ECON. REV. 87 (1999). See also Saul Levmore, *Reasoning and Precedents in Appellate Courts* (2022 available from author).

have reached the same result. One can imagine a legal system that assigned greater precedential value to a 3-0 decision than to a 2-1 decision, but most systems do not do so explicitly. In such a system we would find a different mix of cases, and opinions would be written differently.

With this in mind, consider the situation where an appellate court affirms a decision, and where we can apply the usual assumption that the lower court was more likely than not to get things right. As we will see, this assumption does not much affect the arguments that follow, as they will work whether the lower court judge is often wrong or rarely wrong. Imagine first that the lower court judge is 70% likely to be correct, a number chosen here (not just because of typical reversal rates but, also) because if the losing lawyer observed that this first judge is almost always right, or always affirmed, the lawyer would be less likely to take the case on appeal.¹⁶ In a case where plaintiff loses below, there may be some uncertainty about the size of the judgment that would eventually be paid if the lower court opinion were to be reversed, and the case sent back to the lower court. This adds complexity and it is set aside here, because the main results do not much depend on it. Still, it is worth seeing that if the lower court judge's chance of being right is absurdly low, say 30%, and this is well known (perhaps by the previous rate of reversals or the kind of case involved), then it is likely that an appeal will be filed. But even in this odd case, one appellate judge, who gets things right 70% of the time, is likely to do the job. The opinion below will usually be overturned, and a single appellate judge will set things straight without the need for a panel of three.

But a single *appellate* judge may instead, or also, be of very low quality. That possibility does not do much to strengthen the case for a panel of three, because a below-average circuit court judge, and especially one who is inferior to the lower court judge, might just as likely be the deciding vote on a familiar 2-1 panel. In any event, this is a minor concern, and it is noted here in order to emphasize that the numbers chosen do not much matter.

Imagine, again and perhaps realistically, that the lower court judge is correct with a 70% probability, and that each circuit court judge is somewhat better, with an 80% chance of getting things right.¹⁷ Finally, to make this easier, let us specify that the judges' chances are independent

¹⁶ If it were widely thought that the lower court was just 50% likely to be correct, the parties would normally settle the case without incurring the cost of appellate review. The same might be said for any percentage attached to this first judge, but the inclination not to settle is most easily understood as attached to cases where the opposing litigants have quite different views of the chance of success. See John J. Donohue III, *Opting for the British Rule: Or, if Posner and Shavell Can't Remember the Coase Theorem, Who Will?* 104 HARV. L. REV. 1093 (1991).

¹⁷ The numbers used here are designed to make the arguments clear. In the federal courts, about 10% of appealed cases are reversed, though the numbers differ by subject matter and by what counts as true reversals. See *Just the Facts: U.S. Courts of Appeals*, FEDERAL JUDICIARY, (Dec. 20, 2016), <https://www.uscourts.gov/news/2016/12/20/just-facts-us-courts-appeals>; Theodore Eisenberg, *Appeal Rates and Outcomes in Tried and Nontried Cases: Further Exploration of Anti-Plaintiff Appellate Outcomes*, 1 J. EMPIRICAL LEGAL STUD., 659 (2008) For a careful study of a large set of cases and appeals in state cases, see Theodore Eisenberg & Michael Heise, *Plaintiphobia in State Courts? An Empirical Study*

of one another. The critical argument in this Part is that law should recognize the fact that the lower court judge is also a valued decision-maker, so that when a question is decided by three appellate judges, these are really the second, third, and fourth judges to consider the matter – albeit judges who are probably more likely to be right than judge #1 in the lower court. Returning to medicine, the family practitioner or generalist may be less trained than the expert who is consulted once a problem is identified, but in law, the lower court judge is often as expert as an appellate judge assigned to a case. (If the appellate judge is no more likely, or even less likely, to be correct, the argument here is strengthened rather than weakened.) If the appellate court is 2-1 for reversing the lower court, it is sensible to reverse, because two judges of 80% quality is better than two judges who are 70% and 80%, respectively, likely to be correct. But the value of bringing in three judges to check up on #1 in the lower court is much lower than that of bringing in two medical specialists after the first specialist is consulted. If, on appeal, judges #2 and #3 agree with the lower court judge, #1, there is no gain from involving #4. And if #2 and #3 both want to reverse #1, there is also no point in using #4. The value of the familiar three-judge panel, rather than a single appellate judge, is that the panel might reverse #1 by a 2-1 vote. If the first two appellate judges are split, then #4 is of some value because of the wisdom-of-crowds but, as we have seen, this #4 is no more likely to be correct than was #2. In short, if #2 agrees with the lower court's #1, it is possible but unlikely that something is gained by deploying additional appellate judges. If #3 and #4 disagree with the two preceding judges, there is a small chance of improvement; #1 and #2 might both have been incorrect, with probability of $.3 \times .2 = .06$, but this is mostly offset by the $.04$ chance ($.2 \times .2$) that #3 and #4 are both wrong. With these numbers, there is just a very small chance that a 2-2 decision (including the lower court judge, #1) improves upon the much less expensive 2-0 decision. Moreover, it will very often occur that the appellate judges are unanimous in upholding #1's decision, in which case it is of no consequence to bring in #3 and #4 and, as it turns out, #2 also added nothing. There may be value in solidifying a new precedent, rather than just increasing the chance of what is here called correctness, but this will often be offset by a closer vote or by the uncertainty that #4 brings to the decision. Most objections to this argument will take the form of pointing to the value of judges talking to one another, a feature that Condorcet asked us to avoid, but note that nothing in the patterns offered here prevents a judge from talking to another or even to the lower court judge. This and other objections are taken up presently.

To summarize, if we think of law as looking to reach correct decisions, and we take account of the skill of the lower court judge, the value of making room for a 2-1 appellate decision that reverses the lower court judge must be very small, and perhaps infinitesimal. This leads to the suggestion

of State Court Trials on Appeal, 38 J. LEGAL STUD. 121 (2009) (focusing on the higher rate of reversals from defendant appeals after jury verdicts than from bench trials and plaintiff appeals, and showing that state court reversal rates are higher than federal court reversal rates, and in the neighborhood of 30%). Reversals do not, of course, mean that the previous court was wrong, in the normal sense of the word, a matter discussed in Part IV, but at least the numbers used in the text's several examples are not wildly off the mark. Finally, the rate of affirmance may just mean that appellate courts are deferential rather than really indicating that the lower court was "correct."

that cases ought to be sent on appeal to just one judge, and if that judge decides to affirm the lower court judge, the process might sensibly end there.¹⁸ If the two judges agree as to the outcome of the case, but not as to the *reason* for this outcome, it might make sense to bring in one more judge in order to help establish the precedential value of the decision. But this added wrinkle of judges' agreeing on a result but not on the reasons for it, is a subject taken up elsewhere and set aside here except to say that while there is some wisdom of a crowd, there is also some benefit to clarity, and only a single (appellate) judge can be relied on to provide clarity with a single reason that establishes precedent.¹⁹ In sum, a second reason to prefer one judge over three, despite the modest wisdom-of-crowds feature of the three, is that a single judge avoids the problems faced when combining reasons, or even preferences. Similarly, in our personal lives, it is obvious that asking two friends for advice, rather than one, might add more confusion than clarity.

III. Other Advantages of Multi-Member Appellate Panels

A. The value of discussion and compromise

Condorcet advanced a mathematical reason to encourage or even imagine independence among voters, and this suggests that judges and jurors vote without consulting one another. The wisdom of the crowd is said to be in the aggregation of individual judgments, where each individual is more likely than not to be right (or at least not more likely than not to be wrong). But most appellate judges would say that the ability to interact with one another leads to better decision-making than

¹⁸ It is tempting to ask whether, if the lower court judge's correctness is valuable, the system might be better off investing in two lower court judges, and then limiting appeals to cases where the two disagree, or simply relying on the fact that litigants will be less likely to appeal after a 2-0 lower court opinion. The rule might be that in the case of a 1-1 lower court, the more senior of the two judges prevails. The senior and second lower court judge can simply be regarded as the first appellate judge so that the system would be equivalent to one with an automatic appeal to a single judge, though the first "appeal" is more time consuming and costly. It will raise the probability of a correct decision by increasing the number of voters from one to two, and especially so if there is positive value to discussion and teamwork, but it does so by imposing the costs associated with an automatic appeal, along with the costs of two lower court judges, who must manage to agree or defer to the senior judge regarding various rulings in the course of the trial.

Our current practice in criminal cases could not be more different from the text's argument in favor of a single reviewer. In the case of collateral attacks on criminal judgments by a federal court, the lower court decision is final unless a judge issues a certificate of appealability – indicating that there is a substantial constitutional issue at stake. 28 U.S.C. §2553. If the lower court judge declines to trigger this "appeal," the matter then normally goes to two other judges in sequence, with no communication between them. If the first of these judges issues a certificate, thus reversing the first judge, the question of significance ends there, and the case goes to a panel of three for a decision on the merits. If the first of the two judges declines to certify, then the second considers its importance and that judge can then trigger the panel of three on the merits. In short, the procedure can involve as many as 6 judges, and reflects a view of efficiency that is quite different from that reflected in the text.

¹⁹ See Saul Levmore, *Reasoning and Precedents in Appellate Courts* (2022 available from author).

would separate consideration and blind ballots. We learn from one another as teammates, identifying errors while changing our opinions because of good arguments and observations that others bring to the table. On the other hand, a significant literature warns of group-think as well as the tendency to submit to a bully or a charismatic figure.²⁰ Participants may actually be more likely to hear one another out when a 3-2 divide appears early on, than when a group begins with 5-0 perspective. Either way, as readers with committee experience know, there is a tendency to compromise or be agreeable, until the committee emerges with a unanimous vote. A hard-working committee knows that the larger group, or principal, to whom they report, is far more likely to accept the committee decision if there is not a significant difference of opinion among committee members. Judicial panels share many of these features. The agents want their hard work to be of value. Moreover, unless one is regularly in the minority, it is advantageous to have a norm that a vote within the group will often be reported as a unanimous decision.

Condorcet's theorem does not claim that consultation necessarily makes things worse; it leaves open the possibility that a very skilled individual might be better than a group, and also that a group might benefit from an exchange of views. This is especially so when the group is not very large. Similarly, judges on a panel reviewing a tax case are likely to defer to a colleague who is an expert in that subject. Only a simplistic reading of Condorcet suggests that jurors should be isolated from one another in order to make sure that they vote independently. There is probably no universal answer to the question of whether to choose independence or consultation, and especially so when the group is small. An auto repair shop might be more reliable when kept unaware of an earlier assessment by another shop, while judges may be more likely to reach the correct answer through consultation. We learn from and yet are swayed by one another.

There is an inherent paradox in this matter, and it arises in jury deliberations as well as in committee work, a topic discussed further in Part V. The puzzle is that a strict belief in the wisdom of crowds has voters behaving independently, and yet when a given, contemplative voter sees that a majority of other voters leans one way with respect to a question with a correct answer, the wisdom-of-crowd logic suggests that even the informed, contemplative voter should now abandon her own inclination and go along with the majority of the group, for it is likely to be correct. This might be a reason to insulate voters from one another, because if too many voters follow this logic the wisdom of the crowd will be lost. This is especially the case when votes are taken one-by-one around a room, as later voters are inclined to be influenced – and logically so – by their predecessors. It may also be a good reason for the beyond-a-reasonable-doubt and unanimity requirements governing many juries. A juror who is still thinking about the evidence, and now observes 10 or 11 voters preparing to convict the defendant, or find in a civil case that a defendant was negligent, may sometimes go along with the overwhelming majority because of social

²⁰ For an interesting discussion of the literature, and law's response to it, see Roland Bénabou, *Groupthink in Organizations and Markets*, 80 REV. ECON. STUD., 429, 431 (2013) (using a model to explain that the reason societies set up mechanisms like free speech and whistle-blower protection is because groups would want to "ignore or 'kill' the messengers of bad news").

pressure or some intuition about the wisdom of crowds. A secret ballot may best elicit the true wisdom of the crowd, but it is also more likely to lead to a hung jury.²¹ Juries may be inclined to compromise in order to reach unanimity or to enjoy the community experience, but this is not something anticipated by Condorcet, who dealt with bilateral decisions (guilt or innocence). In any event, the relationship between compromise and the wisdom of crowds is discussed further in Part V in the context of committee work, where the discussion considers the likelihood that the dissenting voter is understood to be more confident than her teammates.

B. The significance of the assigned probability of being correct

The examples offered thus far appeared to depend on the probabilities (of correctness) assigned to the first and then to any subsequent reviewers. A conventional objection to the idea of a single appellate judge is that the ubiquity of multiple appellate panels suggests that they are the product of an efficient evolutionary process. This is consistent with the idea of the wisdom of crowds; we want to be more certain that an appellate court gets things right because it is assigned the power to set precedents, and precedents are themselves efficient because they help avoid the cost of repeatedly inventing the wheel. The current system puts judges with superior skills on these courts and then uses panels of 3, and eventually one of 9, to raise yet further the chances of correctness. But as we have seen, 2-1 is only better than 1-0 when viewed at the start; once we have two judges who disagree, the third is no more likely to be correct than the first. And then once we take the lower court judge into account, very little is gained by having a third and fourth judge.

Experimentation with different probabilities advances the case for a single reviewer in the first round of appeal. Begin with the unlikely case in which it is the lower court judge who is most likely to be right as to the outcome of a case, so that the reason for the familiar three-judge panel must be to tap into the wisdom of the crowd. If the first opinion comes from someone who is very good, then a single reviewer certainly can make things worse rather than better. In the extreme, imagine a case where the first judge is 90% likely to be correct, and then the second opinion is requested from someone who is only 40% likely to be correct. We can continue to assume that #3 and #4 (if called upon) are typical, and correct with .8 probability, a number previously assigned to all appellate judges. The first two judges, #1 and #2, will often agree, and this is the case whether #1 was in fact right or wrong. When they are both correct, the judge with correctness of .9 is followed by .4. When they are both incorrect, .1 is followed by .6. In neither case will additional judges be called upon. With these rather stark numbers, there is a 42% chance (at least when there are appeals) that we will use just two judges and very often (36/42) that will lead to the correct result. But it will often happen (58%) with these numbers that #1, the lower court judge, and #2 disagree, with #2 trying to reverse #1. Most of the time (.9 x .6 compared to .1 x .4) this is

²¹ See Robert F. Holland, *Improving Criminal Jury Verdicts: Learning from the Court-Martial*, 97 J. CRIM. LAW 101, 120 (2006) (discussing the relationship between secret ballots and hung juries, and then suggesting that this hung jury “problem” is avoided in US military trials in part because of a supermajority, rather than unanimity, requirement for conviction).

unfortunate. In these cases, we might have a rule that the matter now goes to #3 or, coming close to the long-standing practice of a panel of 3, to #3 and #4. Infrequently, #1 was wrong and #2 actually improved things (.1 x .4), and we need only fear the unlikely event that #3 and #4 are both wrong (.2 x .2). The more attractive and more plausible sequence is where #1 was correct, and now #3 and #4 will undo #2's mistake on appeal. But with these numbers it will still happen fairly often that at least one of #3 and #4 will have a bad day, as they are *both* right .8 x .8 of the time.²² With a 2-2 decision, #1's correct decision is overruled because the lower court judge is assumed – incorrectly in this case – to be less “accurate.” The court in which #1 sits is even referred to as the “inferior court.” This example is meant mostly for amusement, because we like to think that judges, like doctors, are more likely to be right than wrong, so that the case of the hapless #2 can be ignored. Still, if the first judge is excellent and the appellate judges are only competent, there is room for trouble. If, for example, #1 is 90% likely to be right, and the higher court's judges are (each) only 60% likely to be right, then we need a 3-0 decision to have confidence that we have improved things when the appellate court overturns #1, or we need a majority decision from a larger appellate court. This is fairly intuitive. If one starts out by going to the world's expert to seek help with a troubling medical condition, it is unlikely that getting a second opinion is worthwhile. The basic idea is by now familiar. The more the first opinion is likely to be correct, the less value there is in a second opinion. And if #2 agrees with a competent #1, it is very unlikely that bringing on more reviewers is worthwhile, and especially so if there is no reason to think that #3 and #4 will be much superior to #2. We can already intuit that the reason not to stop (in some cases) after #3 is that the wisdom of crowds might be important in contradicting the expert #2.

On the other hand, if it is #1 who is often wrong, as might be the case if lower court judges are expert in managing witnesses and juries but not in applying or interpreting the law, then a second opinion is surely a good idea. Judge #2 will often, and correctly, overrule #1. But here too, if #2 agrees with the lower court's #1, it is not terribly likely that bringing on more reviewers is worthwhile. First, #3 and #4 may disagree with one another, as there is a decent chance that one of them will be wrong. If they disagree, we might as well have stopped with #2. Alternatively, #3 and #4 may agree, but their input is only worthwhile where #1 was correct and #2 was incorrect, and this confluence of events is quite unlikely. The key idea is to value #1's judgment and to see how unlikely it is that a skilled #1 and a yet more skilled #2 will be *correctly* overcome by #3 and #4. The resources currently spent on panels of three are better spent elsewhere, or even on getting more cases to appellate courts (beginning with just one judge), and then, at most, only when that judge reverses #1 should the matter be sent on to additional appellate judges, #3 and #4. In sum, the argument against the current system of three judges working each appeal is not terribly sensitive to the assumptions made about the likelihood that a given judge in the chain is correct.²³

²² The numbers do matter here, and the argument in the text is yet stronger if the appellate judges are correct with .7 probability, rather than .8. They will then both be right less than half of the time.

²³ The argument would be stronger if the cost of each review were specified. It is, however, hard to estimate these costs, as well as their benefits. There is the cost of each judge's time, the litigation costs for the parties,

A more radical proposal is that things should never go past #2, but this would depend on the probability that #2 is better than #1. This is like saying that when a generalist recommends an expert, the patient might be comfortable relying on this single expert who is, after all, not really reviewing a previous decisionmaker.²⁴

There is added complexity if the putative appellate judges do not share a probability of correctness. The argument was that #2 was (most often) randomly chosen from the pool of appellate judges, so that it is fair to think about the average judge. But perhaps the reason for a panel, as well as the initial opportunity for appeal from judge #1, is that the system recognizes that there are some lesser judges who all too often arrive at the wrong answer to a legal question. The discussion here tries to take account of this in the example where #1 was excellent and #2 inferior, but perhaps the system is designed for the case where #1 frequently errs, and then #2 alone would do little good because #2 also happens to be an inferior judge. One response to this defense of appellate panels of three, is that the better way to deal with a bad employee, not to mention two error-prone employees operating in tandem, is not to bring on more employees but to keep track of employee performance and encourage the poor performers to exit, even if they have life tenure. In any event, the suggestion here works reasonably well even if #2 alone is often wrong. That judge will often disagree with #1, and we can bring in an additional judge, setting aside the most radical version of the argument advanced here.

C. Why not a panel of two?

The central claim here has been that we should start out with a single appellate judge, in large part by taking into account the likelihood that the lower court judge is correct. But if it is too radical to

and the benefit of a better precedent for the future – which is probably the most difficult thing to estimate. Moreover, much depends on what is at stake in the given case. An important constitutional case or a private law case with an enormous amount at stake, might make a small increase in the likelihood of finding the correct answer worthwhile. The text slides pass these points both because the costs and benefits are difficult to estimate and because some of these are surely considered when a yet higher, or *en banc*, group decides to hear a case.

²⁴ A more complete conclusion is that (1) One should go to a specialist before agreeing to a risky procedure. (2) If that doctor suggests a risky, serious procedure, it pays to go for an additional, or second, opinion only if: (a) the second doctor is more likely to be right than the first, or (b) it will make the patient feel better when the two doctors agree. It is, however, hard to put a value on emotional comfort, and of course the cost of greater comfort reflects an insensitivity to the overall costs of health care and insurance, as well as the increase in stress if the two doctors disagree. Finally, (c), it also pays to go to a second specialist if one has the time and resources to go for a third opinion if #1 and #2 disagree. The advice should not be "get a second opinion," but rather "get a second and third opinion unless the second opinion agrees with the first one that you have already received." By way of review, the reason to go for a second opinion is that it gives one a chance to go for a third opinion, and three are more likely to be right than one. In the case of law, however, a panel of three along with a lower court of one, means that we have four experienced minds at work. It only pays to bring on a third judge if #2 disagrees with #1 or, when #1 and #2 agree, then if *both* #3 and #4 are correct and more likely to be right than the first two judges. This is a weak defense of the three-judge appellate panel.

stop after one appellate judge, why not save resources by limiting further review to the use of #3, after #2 tries to reverse the lower court judge, #1. As we have seen, the major reason to engage #3, and even the conventional #4, is to overwhelm #2 when that judge disagrees with #1. The more we think that appellate judges are more likely to be right than lower court judges, the more we might as well stop after #2. If the power of precedent is at stake, so that it is worthwhile to bring on #3 when #1 and #2 disagree, then why bother proceeding to yet another appellate judge. Perhaps a panel of two appellate judges is a good compromise between one and three.²⁵

An argument against this partial retreat from the central argument of this Essay, is the recognition that cases are brought on appeal not from a randomly selected sample of lower court opinions, but precisely when a lawyer thinks that #1 was incorrect.²⁶ The lawyer can also be seen as an informed voter. If appeals are in this way brought when #1, the lower court judge, is more likely to be wrong than is usually the case, then relying on #2 alone to review #1 is attractive. If #2 affirms the lower court, we are done. If #2 reverses the lower court, this Essay has suggested that #3 serve as a tie-breaker.

A better reason to prefer an additional reviewer, in this case #3, is to take account of the likelihood (nearly rejected earlier) that interaction improves the chance of a group's reaching correct decisions. When #1 and #2 disagree, the question is apparently a difficult one, and in just such a case we might want not only the wisdom of a crowd, but also the value of discussion that #2 and #3 would bring to the table. If the two agree, there is also an improved chance of creating a lasting precedent.

IV. "Correct" Decisions

A great deal of the analysis here builds on the assumption that there is a correct answer sought by judges, much as doctors are asked whether to proceed with a serious medical intervention. In some areas of law this is a questionable assumption. Law may be intensely political, so that judicial decisions are a matter of political preferences rather than correct answers. If, on the other hand, judges are out to determine the will of an ambiguous legislature, past or present, or the will of the politicians who nominated and approved their appointment, or even what is good for society, there is likely to be a correct answer, even if it is hard (ever) to know whether a judicial decision was correct. Similarly, if "correct" should refer to social welfare or ethical values, it is again difficult to know when a judge is correct, though there might well be a correct answer. This is much more complicated if different judges have different notions of what is correct, even if they are determined not to let their own preferences get in the way of the agreed upon target.

²⁵ An alternative and less dramatic possibility would be to follow the earlier logic and bring in #4 only when #2 and #3 disagree, after #2 disagreed with #1.

²⁶ This idea was introduced by Steven Shavell, *supra* note 2, and used to explain the different investments in lower and higher courts.

Fortunately, most of the arguments in this Essay do not much depend on how we think of correctness. The most useful approach may be the somewhat circular one: A judge is correct if further review, whether by a higher court, a comparable court, or by others on a panel, affirms the judge's decision. A judge who regularly tries to reverse #1, and is then over-ridden by judges #3 and #4, can be thought of as having erred, if only because of the wisdom of crowds. Similarly, a judge who is regularly reversed by a higher court or an *en banc* review has not made correct decisions.²⁷ This judge imposes substantial costs on the system. One practical lesson is that we might pay more attention to reversal rates, which are already well-known, and encourage judges to retire or to change their thinking if they are much more frequently incorrect (defined this way) than are other judges.²⁸ An easier, and certainly more up-to-date-teaching, approach is to praise the judges who are most frequently affirmed.

A reasonable objection to the approach taken here is that preferences, and especially politics, play a more important role than the pursuit of something called correctness. This is especially the case if we think of judges as devoted to their political preferences. This view is supported by evidence that a judge appointed by one political party behaves differently when placed on a panel with two other judges associated with the same party.²⁹ In constitutional interpretation cases, this may certainly be true. Similarly, some judges are thought to have a taste for or against large corporations, agencies, unions, prosecutors, and environmental concerns. A judge with one of these habits might be identified by the deference given to these entities or concerns. The good news is that this does little damage to the arguments advanced here. Most of these can be refashioned by substituting preferences for correctness. For example, #1 may love big business and so it is sensible to take a business case on appeal from #1's decision if the lower court disappoints one's client. The lawyer hopes to get a #2 with different preferences. But if #1 and #2 agree, there is nothing gained by going to #3 unless the judiciary is known to be stocked with the set of preferences the lawyer seeks. But from the legal system's perspective, and with a roughly equal mix of preferences, if the first two judges disagree, perhaps because they have different preferences, then bringing on #3 (or #3 and #4) adds little because the deciding vote is just as fallible as was #1's in the first place. On the other hand, a 3-0 vote is probably reassuring, because as we approach unanimity with a growing number of judges, it is less likely that pure preferences determined the result, so long as the composition of the judiciary reflects some mix of preferences. This observation returns us to the idea, or passing suggestion, that there is something to be said for

²⁷ Shavell is skeptical that rates of reversal mean much, in part because parties can settle cases before appeal, and they do so with some information about the likelihood that the lower court was incorrect. *See* Shavell, *supra* note 2, at 414.

²⁸ Information on appeals outcomes for district and circuit court judges is compiled by WestLaw. Recall, however, the likelihood that lawyers filter out cases where an affirmation is likely.

²⁹ *See* Eric A. Posner, *Does Political Bias in the Judiciary Matter?: Implications of Judicial Bias Studies for Legal and Constitutional Reform*, 75 U. CHI. L. REV. 853, 854 (2008) (exploring earlier findings about the results from politically mixed panels).

assigning more precedential value to 3-0 decisions. Part V offers something of a contrary argument.

An irony here is that the case for including the lower court judge's decision is now even stronger than before, because while circuit court judges (and certainly Supreme Court justices) appear to be nominated with politics and preferences in mind, this is less so for trial court judges.³⁰ In any event, many lawyers will agree that while constitutional cases attract much attention, court dockets are not filled with constitutional and antitrust cases, where we might also think preferences play a greater role than correctness. Moreover, if we define "correct" to include the likelihood that the vote will not be overturned by other judges, then again most of the arguments here remain firm.

V. Committees, Boards, and Juries

A. Delegation to committees

Committee work in universities, in law firms, in most business settings, and perhaps even in Congress, saves time, and encourages committee members to invest effort because there is less of a collective action problem than if a larger group moves ahead without any delegation to its agents. A committee of three or five might be useful both when a correct answer is sought and when a larger group's preferences are amalgamated. As already noted, but not proved or examined empirically here, it is probably the case that a unanimous committee is less likely to be second guessed than is a divided one. Indeed, in private settings it is often the case that committees do not have an odd number of members, not only because the chair-person has a decisive vote (in formal settings, the chair usually has a deliberative vote and then also a tie-breaking vote, or casting ballot, as it is called) but also because the committee is expected to reach something of a consensus, even if it does so by compromising. In turn, the committee members will know that their work is more valuable if they come to an agreement. In legislatures, committees are often divided along party lines, but in many other settings casual observation suggests that unanimity is common.

Unanimity might come about through compromise, inasmuch as the members know that compromise is the easy path to unanimity and therefore time not wasted. It might also come about through an intuition about the wisdom of crowds. If J sees that her colleagues, K and L, favor giving an academic prize to *A* rather than to *B* or to J's first choice, *C*, J is likely to join her colleagues in order to produce a unanimous decision in the interest of collegiality, and also because J recognizes that her judgment might be idiosyncratic, perhaps because she favors a certain kind

³⁰ The intuition in the text is difficult to support but it is noteworthy that political reliability appears to play a role in the elevation of judges to the circuit courts. Elisha Savcheck et al., *Taking It to the Next Level: The Elevation of District Court Judges to the U.S. Courts of Appeals*, 50 AM. J. POL. SCI. 478 (2006) (demonstrating that district court judges' adherence to the policy preferences of a president is one of the strongest predictors of their being elevated to the circuit level); Stephen Choi et al., *The Role of Competence in Promotions from the Lower Federal Courts*, 44 J. LEGAL STUD. S107 (2015) (demonstrating that competence is not a strong predictor of promotion).

of work or a style of writing. If there are two prizes to be awarded, the committee is likely, when it reports a unanimous vote, to give the more important prize to the choice of the majority, and then to give the secondary honor to the favorite candidate of the minority on the committee. Condorcet might prefer that the majority choose both the first and second place winners, perhaps *A* and *B* in this example, but experience suggests that the quest for unanimity and compromise drives the results to *A* and *C*.

At the risk of moving yet further away from a scientific approach, and introducing more speculation, it is plausible that the quest for unanimity through compromise makes some sense. If the *JKL* committee is divided 2-1, with *L* in dissent, and *L* does not quickly bow to the wisdom of the crowd, not to mention collegiality, *J* and *K* might reason that *L* must be confident of his views. If in fact *J* and *K* is each likely to make the correct decision with .7 probability, and *L* thinks that in this case he is likely to be correct with .95 probability, then *L* ought to prevail, as *JK* (if independently voting) are likely to be wrong with .09 probability (.3 x .3), while *L* is wrong with just .05 probability. Of course, the parties cannot be sure how likely they are to be correct in assessing their own or others' skills, but it might be sensible to compromise rather than to let *JK* completely carry the day.

Put differently, if *L* is in the minority but has no reason to think he is much more likely to be correct than *J* and *K*, then it is reasonable for him to go along with the wisdom of the crowd, small as it is. This is especially the case if he knows that by not doing so he will create more work for the Dean, the Board of Directors, or other principals that asked for a recommendation. It follows that if *J* and *K* know from experience that *L* acts in good faith, and still *L* dissents, they now have reason to think that *L* is fairly sure that he is more likely to be correct than are *J* and *K*. In turn, when the Dean or other principal, or an appellate panel, is faced with a split decision, there is reason to think that *L* is fairly certain but also that *J* and *K* are not convinced that *L*'s assessment of their judgment is accurate. It makes sense in this case to take the 2-1 decision less seriously than a 3-0 decision, and it does explain why committees and corporate boards strive for unanimity. In some settings, split opinions are sent to a larger group for more investigation. In law, we see *en banc* hearings and Supreme Court review of circuit splits in just such situations. The latter has no reason to be unanimous as it will not be reviewed, and the former might contain dissenters both because judges tend to stick to their former opinions and because they might welcome further review. It is interesting and probably sensible that in these further appeals, great weight is not often given to the earlier decision by a slight majority. There is, to be sure, a selection bias here because as we go up the ladder, cases are normally selected for review by the judges. Still, the level-two appellate team, such as the *en banc* group, intuitively that the wisdom of crowds favors the majority in a 2-1 decision preceding its work, but it also knows that the dissenter must be quite confident, for the usual inclination is to compromise in the earlier round or to give in at that first opportunity and accept the wisdom of the crowd.

B. Jury decisions

Juries are only somewhat like committees. The common supermajority requirement encourages each juror to pay attention and to see that his or her opinion might be decisive. Corporate boards and most committees involve repeat players who have reason to work toward a consensus and to trust one another. The supermajority jury may encourage similar behavior among one-off participants. The supermajority requirement, especially for the most serious cases, also makes it more likely that the group will find, or be seen as finding, the correct answer.

An interesting thought experiment involves converting the familiar jury into something that offers a second opinion.³¹ Imagine the choice between a jury of 12, that must reach a unanimous decision, and a scheme that divides the group into two sub-juries of 6. In a criminal case, conviction would require that both sub-juries choose to convict by a 6-0 vote. Would the defendant prefer this arrangement over the familiar one that requires a 12-0 vote? The sub-groups of 6 would probably encourage more participation by each juror. A defendant might not like the sub-group approach because a single juror who is thinking of fighting the group and holding out against a conviction, might give in and figure that if her contrarian reasonable-doubt is justified, it will be emerge in the second sub-group, now considering the case in another room. The single juror need not feel that the weight of the decision is on her alone. On the other hand, a defendant might prefer the sub-group approach, because it is more difficult to stand up against 11 peers than to do so against 5. Jury size does not seem to greatly affect the likelihood of a hung jury,³² but of course the idea here involves two sub-juries, so there is the complication that at least one of them will fail to reach a decision. These are important questions but they do not carry over easily into the design of appellate review, because judges are repeat players and can see whether their decisions are followed by later and higher courts.

Conclusion

³¹ The idea here is not entirely a thought experiment because multiple juries have been understood as an option where there are multiple defendants in the same or similar cases. Judge Posner in *In re High Fructose Corn Syrup Antitrust Litigation*, 361 F.3d 439 (7th Cir. 2004) wrote:

“The power to [empanel multiple juries] in a civil case, if there is such power, has, to our knowledge, been exercised only once, in *Martin v. Bell Helicopter Co.*, 85 F.R.D. 654 (D.Colo.1980). But the existence of the power has not been denied, and we cannot see what there is to bar it. No rule, principle, precedent, statute, regulation, or other source of limitations on the power of district judges stands athwart the procedure that the judge would like to employ. Fundamental to a judge's role as the presiding officer at jury trials is the administration of the rules of evidence in a way that will minimize the likelihood that the jury's verdict will be a product of confusion or inappropriate emotion. Imaginative procedures for averting jury error, as long as they do not violate any legal norm, are to be encouraged rather than discouraged. The procedure the able and experienced district judge wants to employ is orthodox in criminal cases; we cannot see why it should be unacceptable in appropriate civil cases.”

³² Barbara Luppi & Francesco Parisi, *Jury Size and the Hung-Jury Paradox*, 42 J. LEGAL. STUD. 399.

There are many moving parts, not to mention debatable assumptions, in the analysis offered here. The strongest conclusion concerns the structure of appellate review, and the conclusion is often strengthened when the assumptions are tested. Unless we think that lower-court judges are very often wrong and that there are a significant number of appellate judges who would also, with regularity, reach the wrong decision if asked for a second opinion, it is probably a waste of resources to have appeals go to a panel of three judges. It would be better to have appeals go to a single judge. It might (or might not) be worthwhile to bring on additional appellate judges if the first two judges disagree. A compromise is to bring on an additional judge if the first appellate judge chooses to reverse the lower court, or perhaps if there is reason to think that the precedential value of a decision will be enhanced by involving more judges. This conclusion holds even if there is disagreement about the nature of a “correct” answer in law and, at least arguably, even if judges do better when they discuss cases in teams. Most significantly, it draws on, rather than ignores, the wisdom of crowds.