

Appellate Panels and Second Opinions

Saul Levmore*

ABSTRACT

Judges working in panels of three offer something of the “wisdom of crowds,” and especially so when two judges alone would be divided. And yet, in that case, the tie-breaking judge is no more likely to be “correct” than was the first judge. This Article 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 several counterintuitive results. It leads to the conclusion that most appellate processes should be restructured so that one judge alone reviews a lower court. If the single appellate judge disagrees with and seeks to reverse the lower court, this Article suggests that just one more judge should enter the fray, in part because the lower court judge is also to be valued, so that there is a 2-1 decision at hand. The argument uses some probabilistic thinking and therefore has counterparts in other settings, including jury decisions, where the familiar question of when to seek and pay for a second opinion arises. As the argument proceeds, the value of long-lasting rules and other aims are brought into play.

Table of Contents

I. INTRODUCTION	812
II. SECOND OPINIONS QUITE GENERALLY	814
A. <i>Going for a Second Opinion Where There Is a Clear Right Answer</i>	814
B. <i>Why 2-1 Is Better Than 1-0</i>	819
III. SECOND AND THIRD OPINIONS IN THE JUDICIARY	822
IV. OTHER ADVANTAGES OF MULTI-MEMBER APPELLATE PANELS	825
A. <i>The Value of Discussion and Compromise</i>	825
B. <i>The Significance of the Assigned Probability of Being Correct</i> ...	827

* I am grateful for the discussions I enjoyed with Dylan Baker, Frank Easterbrook, Dylan Fane, Eliot Levmore, and workshop participants at Stanford, Columbia, and the University of Chicago and the University of Texas Law Schools.

C. <i>Why Not a Panel of Two?</i>	831
V. “CORRECT” DECISIONS.....	832
VI. COMMITTEES, BOARDS, AND JURIES.....	834
A. <i>Delegation to Committees</i>	834
B. <i>Jury Decisions</i>	836
VII. CONCLUSION	837

I. INTRODUCTION

Appellate review can be understood as an opportunity to correct errors made by lower courts and, because appellate reviews involve more than one judge, 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 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 aiming 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.

These perspectives have counterparts in other settings where the familiar question of when to seek and pay for a second opinion arises. Most second opinions, whether sought before agreeing to a medical procedure or contracting for an auto repair, are given by a single analyst. In contrast, appellate review in the federal and most state systems normally involves three jurists, and possibly more in the event of a

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 William M. Richman & William L. Reynolds, *INJUSTICE ON APPEAL: THE UNITED STATES COURTS OF APPEALS IN CRISIS 1–9* (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 V. When decisions are viewed as correct, precedential value and uniformity become especially attractive.

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

further appeal.³ This Article 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 several counterintuitive results. The most striking conclusion is that 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. However, 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 counterintuitive, but reasonable, conclusions to reach once the logic of appellate review, and second opinions quite generally, is examined. The first appellate judge might decide whether further review is in order. Alternatively, the appellate process could stop after one review—subject to the Supreme Court’s decision to take the case. These and other conclusions are examined here but are 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. As the argument proceeds, the value of long-lasting rules and other aims are brought into play.

Part II 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 opinion is poorly conceived. Part III 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 IV 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 in this Article about the probability that a judge is correct. Part V turns to the possibility that appellate judges are not deployed to find correct answers, perhaps because there is often no such thing, but rather

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 are partly borne by the government and might be overused (or overused). The litigant is not charged for the cost of judging, but the litigant is also not rewarded for the value of creating precedents used by others. The optimal rate of review is not the subject here.

to reflect and aggregate preferences—an idea familiar to readers who regard much of what judges do as reflecting political preferences. Part VI extends the analysis of appellate review to committees, to 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.

II. SECOND OPINIONS QUITE GENERALLY

A. *Going for a Second Opinion Where There Is a Clear Right Answer*

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, acting as a gate-keeper, to move ahead with a serious intervention without consultation. The analysis is not quite the same 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 obviously wise to seek two specialists' opinions. 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, especially so if the first recommends something drastic. Indeed, many health insurance companies encourage, or even require and finance, second opinions in specified circumstances.⁴

There is an equilibrium. When a 90% (probability of being correct) second opinion is 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 first responders who are simply competent.⁵

4. See Robert B. Grant, *Outpatient Surgery: Helping to Contain Health Care Costs*, MONTHLY LAB. REV., Nov. 1992, at 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. However, as argued in this Article, the more we can save costs by having a good lower court, or first doctor, and then a smaller number of reviewers, the more sensible it is to invest in good first responders. An

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, who is 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. Of course, law might have some other goals (a matter explored in Part V) and the definition of "correct" is contestable.

The assumption of correctness in this Article ignores the idea that law might seek to appear legitimate rather than correct. Statistically minded observers will have more confidence in a 3-2 or 2-1 decision than a 1-0 decision. More draws from an urn always improve one's assessment of what is in the urn, so to speak. However, unanimity (even with one "voter") may have value even if it is not the best we can do in terms of finding a right answer. Many citizens might regard a 1-0 decision as more reliable than a 2-1 decision, just as many patients will feel uncomfortable once a respected doctor disagrees with others. This supports the idea of pausing after a 2-0 vote, rather than asking for a third judge's opinion. Nevertheless, the argument here proceeds with the idea that a correct decision is sought and is available. As we will see, the real advantage of a smaller, unanimous vote is that it is more likely to avoid the problem of multiple voters relying on different reasons.

To simplify the analysis, imagine that specialist #1 recommends that a medical 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 intervention will lead to disaster. 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

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 actor knows that errors are likely to be caught at the next step. This is likely to be a problem when the first actor 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.

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 time and money are available to obtain additional opinions before agreeing to an invasive procedure, the well-known “wisdom of crowds,” or Condorcet Jury Theorem, prevails.⁶ 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 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

6. 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. It reaches 1 far more quickly the more there is a supermajority. 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 otherwise followed by them. Without compromise (a topic discussed in Part VI), 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 only a bit lower than the chance of getting 8 or more correct opinions. Even a bare majority is better than a single doctor who is skilled enough to be correct with the 66.7% probability imagined here. It is easy to explain Condorcet (and his followers) nearly brushing aside this difficult topic because he was personally interested in unanimous jury decisions with the death penalty in sight.

For a clear but also critical view of the Condorcet Jury Theorem as applied to law, see generally 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).

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. The same will be true with respect to judicial decision-making, as we will see.

Return then to the value of seeking 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 #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 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 assessment, the patient is unsure how to proceed and chooses instead to flip a coin, the patient is no better off than stopping after #1, and might just as well have proceeded without seeking advice.⁷ 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 the usual assumption of independent judgments.

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

7. This may seem puzzling, inasmuch as the patient has useful information after #1 and then 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 8, is statistically useful but sometimes of no help to the patient.

patient correctly and then by appealing to many readers' intuitions. When each voter, or decision-maker, is more than 50% likely to be right, we should amalgamate *independent* judgments. The preference for independent assessments, or judicial decision-making, is not simply to 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, 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, an 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 should be understood to come along with a requirement that one must be ready to seek a third—unless the second opinion will be acquired from a source significantly more likely to be correct than that which produced the first opinion. If there is no time to get a third opinion then, counterintuitively, it is usually wasteful to solicit a second opinion. In the case of medical intervention, even the idea that second opinions might provide peace of mind is questionable, because 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 involving two doctors who are about equally likely to be correct, the patient is relying on the quality of #3's opinion, and by hypothesis this is identical to 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. In the case 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. A single judge also saves time and money. To be sure, a "correct" decision is likely to be an appealing goal with respect to medical care, but where judging is concerned, the notion of correctness is subject to debate. 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. If the highest court is added to the argument, a correct decision might be defined as one that stands the test of time and is not reversed by a future

court or legislature. It reflects the idea that law improves over time, as there is experience, discussion, and opportunities to review. This notion of correctness is revisited in Part V.

B. Why 2-1 Is Better Than 1-0

The second conclusion just stated (that one should forego the second opinion because after a 1-1 split, the third opinion is no more likely to be right than the first) is mostly wrong. It is wrong in a way that reveals an important point about additional opinions, whether by doctors or judges looking for right answers. 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, 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.⁸ Essentially, the reason to go for

8. There are several ways to see this striking point. If the first doctor is correct, the chance of #1 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 right) \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.

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 the first opinion.⁹

In terms of the numbers used here, if it pays to go to #2 it is because, assuming there will be time for #3, there is a $2/27$ chance of correcting an error by the first decision-maker.¹⁰ 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.¹¹ Still, a second opinion is worth something, but only if it is understood as opening the door for three opinions rather than one. Put differently, decision-makers #2 and

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 have been correct. Competent or exceptional review is most helpful in the middle.

9. 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.

10. 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 8.

11. 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), and the benefit of a correct decision, and then multiply that 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 Article). The familiar 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. However, 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. 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.

#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 a 2-1 decision is better than a 1-0 decision (unless the first decision-maker is much more likely to be correct than the other two) because the 2 may be reversing the incorrect 1, 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 a 5-4 decision reverses a 2-1 decision, we are no better off unless the justices are more likely to be "correct" than the majority of judges below.

There is, however, a catch. 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 of interest not because there is a tiny chance that it is correct compared to 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; our assumption or imagination was off base. It is more likely that many of these 99 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 matter right.¹²

The discussion in this Part has demonstrated that there is some wisdom of the crowd even if a 2-1 decision might simply mean that the third voter, who was no more likely to be right than the first voter, decided things. 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 to return a simple majority vote, and this produces at least a 2-vote margin.¹³ It is easy to see that a 7-5 decision is better than 1-0. However,

12. 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 generally 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 correct answers. On the one hand, "extremist" voters should be happy to follow the wisdom-of-the-crowd but, on the other hand, they may genuinely think that there are issues regarding whether a particular individual is more likely to know the correct answer. The matter is discussed further in Part V.

13. Once we abandon the requirement of unanimity, other issues arise. For example, one majority might vote for liability but a different majority might vote for the particular

it is more difficult to see the point of empaneling an odd number of jurors—such as 11 jurors who might reach a 6-5 result—when a 1-0 decision is almost equally informative. Of course, a significant super-majority decision is more reliable than a 1-0 decision, but requiring a super-majority in civil cases will lead to hung juries or to less reliable compromises within the jury. Moreover, a super-majority requirement puts pressure on the framing of the question presented to the jury. These are matters best left for another day. The present point is simply that a majority of 7-5 is plainly superior to a 1-0 decision.

III. 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 first 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.¹⁴

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—have reached the same result. One can imagine a legal system that assigns 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 even where 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 it matches 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.¹⁵ In a case where the plaintiff loses

determination of damages. See Shari Seidman Diamond, Mary R. Rose & Beth Murphy, *Revisiting the Unanimity Requirement: The Behavior of the Non-Unanimous Civil Jury*, 100 NW. U. L. REV. 201, 210 (2006) (suggesting the possibility of excluding jurors who voted against liability from participating in the vote on damages).

14. See generally Maxwell L. Stearns, *Should Justices Ever Switch Votes?: Miller v. Albright in Social Choice Perspective*, 7 SUP. CT. ECON. REV. 87 (1999); Saul Levmore, *Fractured Majorities and Their Reasons*, 127 PENN ST. L. REV. 331 (2023).

15. 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

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 only to emphasize that the numbers chosen do not much matter.

Imagine, again and perhaps realistically, that the probability that a lower court judge is correct is 70%, and that each circuit court judge is somewhat better, with an 80% chance of getting things right.¹⁶ Finally, to make this easier, let us specify that the judges' chances are independent 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

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 generally* 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).

16. 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 generally* *Just the Facts: U.S. Courts of Appeals*, U.S. CTS. (Dec. 20, 2016), <http://bit.ly/3WXEiNL>; Theodore Eisenberg, *Appeal Rates and Outcomes in Tried and Nontried Cases: Further Exploration of Anti-Plaintiff Appellate Outcomes*, 1 J. EMPIRICAL LEGAL STUD. 659 (2004). For a careful study of a large set of cases and appeals in state cases, *see generally* Theodore Eisenberg & Michael Heise, *Plaintiphobia in State Courts? An Empirical Study 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, 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 V, 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" in the usual sense of the word.

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 much of an 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 are better than two judges who are 70% and 80%, respectively, likely to be correct.

The value of bringing in three judges to check up on #1 in the lower court is small. 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, produced by the lower court and one appellate judge.

It will very often occur that multiple appellate judges are unanimous in upholding #1's decision; in that 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 value 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, but note that nothing in the patterns offered here prevents a judge from talking to another judge 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 is very small—perhaps infinitesimal. This leads to the suggestion 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 should end

there.¹⁷ 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 for another day, and it is 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.¹⁸ 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.

IV. OTHER ADVANTAGES OF MULTI-MEMBER APPELLATE PANELS

A. *The Value of Discussion and Compromise*

The Condorcet Jury Theorem¹⁹ offers a mathematical reason to imagine or even encourage independence among voters, and this suggests that judges and jurors should vote without consulting one

17. 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 than the current system which employs just one lower court judge. Moving to a two-judge lower court will raise the probability of a correct decision by increasing the number of voters from one to two—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. *See* 28 U.S.C. §2253. 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.

18. *See* Levmore, *supra* note 14.

19. *See supra* note 6.

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 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 groupthink 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, so that the group 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 will appear advantageous for the group to reach a unanimous decision.

Consultation does not, of course, necessarily make things worse; the idea is simply that a very skilled individual might be better than a group, and 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 dedicated contrarian would suggest that jurors should be isolated from one another to ensure 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 VI. The puzzle is that a strict belief in the wisdom of the crowd has voters behaving independently. And yet, when a contemplative voter sees that a majority of other voters leans one way with respect to a

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").

question with a correct answer, the wisdom-of-the-crowd logic suggests that even the informed voter should now abandon her own inclination and go along with the majority, for it is likely to be correct. This might be a reason to insulate voters from one another in some settings; 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, in a civil case, find the defendant was negligent, may sometimes go along with the overwhelming majority because of social pressure or an 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. This is not something anticipated in the literature on the wisdom of crowds, which usually deals with bilateral decisions (guilt or innocence) where there is little room for compromise on the most important question. In any event, the relationship between compromise and the wisdom of crowds is discussed further in Part VI, where the discussion considers the likelihood that the dissenting voter on a committee 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 fairly 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 reinventing the wheel. The current system puts judges with superior skills on these courts and then uses panels of three, and eventually one of nine, to raise yet higher 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

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

no more likely to be correct than the first. And when we take the lower court judge into account, very little is gained by involving 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 likely to be correct, 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, creating a 36% chance both are correct. When they are both incorrect, .1 is followed by .6, creating a 6% chance both are incorrect. In neither case will additional judges be called upon because the two are in agreement. The good news is that, when judges #1 and #2 agree, there is a 36/42 percent likelihood that the two will be correct. The bad news is that, with 58% probability, the two will disagree and a third judge will be called upon. It is unfortunate because by assumption this third judge is not as good as the first, lower court judge. In these cases, we would prefer the long-standing practice of a panel of three, that is adding #3 and #4 to the picture, because they will overcome the impact of that poor (.4) judge #2. In other words, rather 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 just .8 x .8 (64%) 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 completeness, 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

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.

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 decision²³ 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. This is 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 experts 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 on getting more cases to appellate courts (beginning with just one judge).²⁴ Only when a single appellate judge (now known as #2) wants to reverse #1 should the matter be sent on to additional appellate judges. In sum, the argument against the current

23. If the three judges share a .4 chance of being wrong, and their judgments are independent, then there is a .064 chance (.4 x .4 x .4) that a 3-0 decision is wrong. The 93.6% chance of being correct is more than the first judge's 90%.

24. 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 and the litigation costs for the parties, and there is 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 involved might make a small increase in the likelihood of finding the correct answer worthwhile. The text slides past 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.

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. A more radical proposal is that the process should never go past #2, but this requires the implausible assumption that virtually every appellate judge is more likely to be correct than almost every lower court (#1) judge. Judges are unlikely to be like doctors; we do think that a single well-recommended specialist is almost always more likely to be correct than the generalist, who is #1 in that chain.²⁵

If the idea of saving resources is unconvincing, another way to think of the argument advanced here is that a second judge is valuable because the first judge might have a bad day or simply be incompetent. One response to this defense of three-judge panels, that they are inexpensive, is to ask why the same logic does not encourage five-judge panels? A better and far more common reaction is to say rather simply that with a great deal at stake, whether in a courtroom or in a hospital, it is worth getting a second opinion just in case the first decision-maker is wrong. A 2-0 decision is then somewhat wasteful, but a 1-1 decision offers the opportunity to correct a disastrous situation created by a mistaken #1. If the first decision-maker is wrong one in ten times, then the likelihood of two mistakes in a row is one in one hundred, and that is very unlikely. We pay a small amount for that difference. If #2 is far less likely to err than #1, we can stop there. But if the question is which one had a bad day, then we need a third judge or doctor. This is simply another way of looking at the argument for a single-judge appellate panel, with another judge or two brought on only if the first appellate judge disagrees with the judge below.

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

25. A more complete conclusion is that: (1) One should go to a specialist before agreeing to a risky procedure, and (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 are more likely to be right than the first two judges. This is a weak defense of the three-judge appellate panel.

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 stop after one appellate judge, why not save resources by limiting further review to the use of #3, after #2 tries to reverse #1, the lower court judge. 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 likely appellate judges are to be right, 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, why bother proceeding to yet another appellate judge? Perhaps a panel of two appellate judges is a good compromise between one and three.²⁶

An argument against this partial retreat from the central argument of this Article 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.²⁷ 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 Article has suggested that #3 should serve as a tiebreaker.

26. 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.

27. See generally Shavell, *supra* note 2 (introducing this idea and using it to explain the different investments in lower and higher courts).

A better reason to prefer an additional reviewer, in this case #3, is to take account of the likelihood that interaction really does improve the chance of a group reaching correct decisions. When #1 and #2 disagree, the question is apparently a difficult one, and in 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.

V. "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 proceeding with a serious medical intervention is the correct thing to do. As noted, the argument proceeds with the assumption that the goal is to find a correct answer rather than to make a population comfortable with an outcome. 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 to (ever) 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.

Fortunately, most of the arguments here 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. An appellate 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.²⁸ This judge imposes substantial costs on the system. One practical lesson is that we might pay yet more attention to reversal rates than is presently the case, and then encourage judges to change their thinking or even to retire if they are much more frequently incorrect

28. 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.

(defined this way) than are other judges.²⁹ An easier approach is to praise the judges who are most frequently affirmed, and to hope that praise influences judicial behavior.

One advantage of defining correctness as a decision that will not be reversed by a higher court, or perhaps even by a legislature, avoids the problem of an apparent majority—glorified by the Condorcet Jury Theorem—reaching its decision by way of inconsistent reasons. Judges #1 and #2 may favor an outcome that is disfavored by #3, but the two judges in the majority may disagree with one another's reasoning. Indeed, where #1 and #3 agree that #2's reasoning is wrong, the wisdom of the crowd is that the outcome favored by #1 and #2 is thus wrong, because at least one of their reasons is wrong. This possibility makes a majority's decision less attractive than it first seems. The relationship between outcomes and reasons for outcomes is addressed in other work, and avoided here by defining correctness not as something that is right, like $2 + 2 = 4$, but rather as something that will not be contradicted by a higher court or, arguably, by a legislature that tells judges they were wrong.³⁰

A reasonable objection to the approach taken here is that preferences, and especially politics, play a more important role in understanding what judges do than the pursuit of something called correctness. This is especially the case if we think of judges as devoted to their political preferences or sponsors. This view is supported by evidence that a judge associated with one political party behaves differently when placed on a panel with two other judges who had been appointed by a president from another party.³¹ 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 organizations. A judge with one of these apparent preferences might be identified by the deference given to various entities (like corporations or specific agencies). The good news is that this does little damage to the arguments advanced here. Most of the arguments here can be refashioned by substituting preferences for correctness. For example, judge #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

29. Information on appellate outcomes for district and circuit court judges is compiled by WestLaw, but the information must be read with an understanding that lawyers filter out cases where an affirmation is likely.

30. See Levmore, *supra* note 14.

31. 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).

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 confidence that the judges found the correct result 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 that there is something to be said for assigning more precedential value to 3-0 decisions. Part VI 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.³² It is apparent that most trial court dockets are not filled with antitrust cases or other matters where judicial decisions might well align with preferences. In most cases "correct" is comfortably defined as secure over time.

VI. COMMITTEES, BOARDS, AND JURIES

A. *Delegation to Committees*

Committee work in universities, law firms, 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, 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

32. 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. *See generally* Elisha Carol Savchak 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 J. Choi, Mitu Gulati & Eric A. Posner, *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).

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 of work or 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. A political scientist, enamored with the Condorcet Jury Theorem and the likelihood that a majority is more likely to be correct than a minority, might prefer that the majority choose both the first and second place winners—much as a majority normally chooses both senators from one state or even the president and vice-president. But committees are not populated with political scientists, and experience suggests that the quest for unanimity and compromise drives the results to *A* and *C*, rather than to *A* and *B*, as both are preferred by members *J* and *K*.

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 *J-K-L* 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. If, in fact, *J* and *K* are 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 *J-K* (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 *J-K* completely carry the day.

In a slightly different scenario, 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 an appellate panel is faced with a split decision, there is reason to think that *L* is fairly certain that she is correct, but also to think 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 because of the recognition—and perhaps also the desire to find—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.³³ 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 six. 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 six 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 emerge in the second sub-group 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 five. Jury size does not seem to greatly affect the likelihood of a hung jury,³⁴ so two (sub-) juries raise the probability of a hung jury. These are important questions, but they do not easily carry over to the design of appellate review, because judges are repeat players and can see whether their decisions are followed by later and higher courts.

VII. CONCLUSION

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

33. 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* 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.

In re High Fructose Corn Syrup Antitrust Litig., 361 F.3d 439, 441 (7th Cir. 2004).

34. See Barbara Luppi & Francesco Parisi, *Jury Size and the Hung-Jury Paradox*, 42 J. LEGAL. STUD. 399, 415 (2013).

number of appellate judges who would be regularly reversed if they were acting alone, there is no good reason to have most appeals go to a panel of three judges. Larger appellate courts waste resources. And if the argument about resources is unconvincing, and yet there is no clamor for five-judge panels, then the case for a single appellate judge is that this is enough to reverse the occasional error by a lower court. Only if these two judges disagree do we need a third or fourth opinion. A compromise is to bring on an additional judge if the first two judges agree that the precedential value of a decision will be enhanced by involving more judges. The case for a single appellate judge 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 the crowd.