ALL OR NOTHING: EXPLAINING THE SIZE OF SUPREME COURT MAJORITIES

PAUL H. EDELMAN* AND SUZANNA SHERRY**

In this Article, Professors Edelman and Sherry use a probabilistic model to explore the process of coalition formation on the United States Supreme Court. They identify coalition formation as a Markov process with absorbing states and examine voting patterns from twelve Court Terms. On the basis of their data, they conclude that Justices are reluctant to remain in small minorities. Surprisingly, however, they also find that a three-Justice minority coalition is less likely to suffer defections than a four-Justice minority coalition. This counterintuitive result suggests that while in general it is minority Justices rather than majority Justices who drive the process of coalition formation, five-Justice majorities may be particularly interested in attracting additional votes. The Article closes with suggestions for future research.

INTRODUCTION

The United States Supreme Court has nine Justices, and it takes five to make a majority. Based on this information alone, what might

* Professor of Mathematics, University of Minnesota <edelman@math.umn.edu>.
** Earl R. Larson Professor of Civil Rights and Civil Liberties Law, University of Minnesota <sherry@umn.edu>. We thank Steven Brams, Jim Chen, Frank Cross, Lee Epstein, Dan Farber, Barry Friedman, and Gordon Silverstein for comments on earlier drafts of this paper.

you suppose are the two most common sizes of a majority coalition? What about the least common?

Previous research has suggested *both* that the larger the majority, the less likely it is to make any effort to attract further votes[1] *and* that the smaller the number of dissenters, the more likely any given dissenter is to defect to the majority.[2] These two propositions create some tension, leaving us unable to predict how likely it is that any given majority might garner additional votes: for example, whether a seven-to-two initial vote eventually will become an eight-to-one or even a unanimous opinion.

Even more intriguing than these tensions in prior research are the actual data. Examining twelve Supreme Court Terms spread over four decades, we found that, in ten of the twelve Terms, unanimous opinions and five-to-four splits were two of the three most common results. In almost half of the Terms, unanimous decisions and five-to-four decisions were *the* two most common outcomes. On the other hand, eight-to-one decisions were among the *least* common. In other words, the Court seems to vacillate between maximum unanimity and maximum division.[3] What explains this pattern?

After a brief discussion in Part I of other scholars' unsuccessful attempts to explain the pattern of majority coalition sizes, we propose in Part II a new model for understanding the distribution of majority coalition sizes. Our model views coalition formation as a stochastic process, with each non-majority Justice continually reacting to the size of the majority coalition at any given time. In particular, we posit a model based on a basic Markov process with absorbing states.[4]

Solving the equations of our model for the twelve Terms leads us to conclude in Part III that, with one extremely interesting caveat, the smaller the number of dissenters, the more likely it is that those dissenters will be persuaded to join the majority, despite the fact that

1. *See* Paul J. Wahlbeck et al., *Marshaling the Court: Bargaining and Accommodation on the United States Supreme Court*, 42 AM. J. POL. SCI. 294, 312 (1998).

2. *See* Saul Brenner et al., *Fluidity and Coalition Sizes on the Supreme Court*, 36 JURIMETRICS 245, 249–52 (1996); Robert H. Dorff & Saul Brenner, *Conformity Voting on the United States Supreme Court*, 54 J. POL. 762, 768 (1992); Forrest Maltzman & Paul J. Wahlbeck, *Strategic Policy Considerations and Voting Fluidity on the Burger Court*, 90 AM. POL. SCI. REV. 581, 587–89 (1996).

3. Aggregating all twelve Terms shows that unanimous opinions are most common (27%), followed by five-to-four opinions (24%). Eight-to-one (14%) and seven-to-two (13%) decisions are the least common. Six-to-three decisions fall in the middle (21%). Because of rounding, these percentages do not total 100%.

4. Both the Markov process and the mathematical basis for solving it are described in Appendix A.

the majority may make no particular effort to attract votes. We call this pattern the "bandwagon effect," and our model allows us not only to validate it but to measure it. Thus, our findings support and elaborate on prior research suggesting that a Justice who is a lone dissenter is the most likely to change his vote.

This conclusion is important not only for its quantification of prior assumptions about voting patterns, but also because it suggests quite strongly that future research should go in particular directions. Our findings demonstrate that, in exploring Supreme Court voting behavior, researchers should focus on dissenters rather than on the majority. Moreover, because the mathematical probabilities *by themselves* sufficiently explain the actual distributions of coalition sizes, our data imply that focusing on either individual Justices or on case-specific outcome preferences is largely irrelevant, except to explain minor fluctuations from Court to Court or Term to Term.

Our second conclusion is perhaps even more startling: the most stable size for a majority opinion is six Justices. In other words, the least likely Justice to join a majority coalition is one who is in a three-Justice minority (*not* a four-Justice minority, as one might expect from the other results). Neither cooperative game theory nor the bandwagon effect would predict this result. We discuss the implications of this finding at the end of Part III.

I. PREVIOUS EXPLANATIONS

Previous explanations of voting patterns are insufficient primarily because they explain only part of the data and conflict with other data. Many theories are offered, for example, to explain the prevalence of unanimous opinions, but all of them would *also* predict a low occurrence of five-to-four opinions and a relatively high occurrence of eight-to-one opinions. Although standard game theory analysis can explain the prevalence of five-to-four opinions, it is completely unable to explain the high occurrence of unanimous outcomes. Indeed, it is the "all-or-nothing" quality of the results that is most difficult to explain. Any successful model must explain consistently *all* of the data—including not only the prevalence of both maximally united and maximally divided Courts, but also the dearth of eight-to-one decisions.

One possible explanation for the existence of so many unanimous cases is that many of the cases before the Court are simply easy calls. But if it is ease of decision that produces unanimous opinions, then why are there so few eight-to-one cases, which are presumably just a little bit less easy? And why do the cases seem to be either very easy (yielding unanimous decisions) or very hard (yielding five-to-four decisions), with fewer cases in between? Moreover, it seems implausible that very many "easy" cases actually reach the Supreme Court.[5]

A second popular explanation for the prevalence of unanimous opinions is similarly flawed. Several scholars have argued that there is considerable institutional pressure to achieve unanimity, especially in important cases.[6] Unanimous decisions are seen as more compelling and less susceptible to either outright defiance or overturning by Congress or a subsequent Court. The classic examples, of course, are *Brown v. Board of Education*[7] and *Cooper v. Aaron*.[8] Both cases were not only unanimous, but signed by all nine Justices, and, in *Cooper*, Justice Frankfurter was persuaded not to release his concurring opinion until a week after the decision came down.[9]

While institutional pressure may explain some individual cases, it does not explain very many of the roughly one-third of all cases that are decided unanimously. The vast majority of these cases are not of such social or political import that the public or the legislature will care whether the outcome is unanimous or not. As an exercise in this regard, think back to the 1997–1998 Term and attempt to recall even one of the thirty-nine unanimous decisions of that Term.[10]

---

5. For another critique of this explanation, see Saul Brenner & Theodore S. Arrington, *Unanimous Decision Making on the U.S. Supreme Court: Case Stimuli and Judicial Attitudes*, 9 POL. BEHAV. 75, 83–84 (1987). A variant on this explanation is a recent model suggesting that more complex cases generate more opinions. *See* Scott P. Johnson, *The Influence of Case Complexity on the Opinion Writing of the Rehnquist Court*, 25 OHIO N.U. L. REV. 45, 58 (1999). Not only does this variant still leave the overall pattern of coalition-size distributions unexplained, but it also considers the cases as falling into only two categories, simple and complex. There is no examination, for example, of the relationship between the *relative* complexity and the number of separate opinions.

6. *See, e.g.*, LEE EPSTEIN & JACK KNIGHT, THE CHOICES JUSTICES MAKE 107 (1997); David W. Rohde, *Policy Goals and Opinion Coalitions in the Supreme Court*, 16 MIDWEST J. POL. SCI. 208, 215–16 (1972); Jeffrey R. Lax, Unanimity and Dissent in the Supreme Court: The Conflict of Consensus 1, 2 (Aug. 1996) (unpublished manuscript, on file with authors).

7. 347 U.S. 483 (1954).

8. 358 U.S. 1 (1958).

9. *Cooper* was issued on September 29, 1958, and Justice Frankfurter did not file his concurring opinion until October 6. *See Cooper*, 358 U.S. at 20 n.* (Frankfurter, J., concurring). On the negotiations leading up to this delay, see Dennis J. Hutchinson, *Unanimity and Desegregation: Decisionmaking in the Supreme Court, 1948–1958*, 68 GEO. L.J. 1, 82–83 (1979).

10. The unanimous decisions included three cases interpreting provisions of the federal bankruptcy act, two interpretations of the statutes governing removal jurisdiction,

Moreover, the "institutional pressure" hypothesis is flatly inconsistent with the prevalence of five-to-four decisions. Again, as with the "easy case" explanation, the problem lies in explaining why the data do not fall into a smooth decreasing curve but rather into two spikes at opposite ends. If we adhere to the "institutional pressure" theory, then the actual data suggest that cases are either very important or not important at all, with few in between.

A third explanation for the frequency of unanimous opinions is a variant on the second: the greater the potential threat that any particular case poses to the Court itself, the larger the majority coalition is likely to be.[11] This theory has been critiqued elsewhere[12] and is suspect because it focuses only on civil liberties cases and only on the Warren Court. Both that Court and that type of case may not be representative. In any case, this theory, like the others, does not explain the actual pattern of distribution unless one concludes that cases most commonly present either a great potential threat (unanimous decisions) or very little potential threat (five-to-four decisions), but rarely anything in between.

Turning to the other side of the data, the large number of five-to-four decisions can be explained using cooperative game theory, which suggests that the most typical coalitions are minimally winning ones—in this context, a five-to-four decision.[13] But again, this does not

---

an ERISA case, a case establishing the statute of limitations for the federal Multiemployer Pension Plan Amendment Act, and a case involving the Death on the High Seas Act (to name some of the cases in which social importance or institutional pressure was unlikely to play a role). *See* Wisconsin Dep't of Corrections v. Schacht, 524 U.S. 381 (1998) (removal); Dooley v. Korean Air Lines Co., 524 U.S. 116 (1998) (Death on the High Seas Act); Geissal v. Moore Med. Co., 524 U.S. 74 (1998) (ERISA); Cohen v. De La Cruz, 523 U.S. 213 (1998) (Bankruptcy Act); Kawaauhau v. Geiger, 523 U.S. 57 (1998) (Bankruptcy Act); Rivet v. Regions Bank of La., 522 U.S. 470 (1998) (removal); Fidelity Fin. Serv. v. Fink, 522 U.S. 211 (1998) (Bankruptcy Act); Bay Area Laundry v. Ferbar Corp. of Cal., 522 U.S. 192 (1997) (MPPAA).

Among the cases decided by a *divided* Court, and therefore of less social or political importance according to the "institutional pressure" theory, were *Clinton v. City of New York*, 524 U.S. 417 (1998) (invalidating the line item veto), *County of Sacramento v. Lewis*, 523 U.S. 833 (1998) (immunizing the police from liability for a death caused by a high speed chase), *Burlington Industries v. Ellerth*, 524 U.S. 742 (1998) (imposing vicarious liability on employers for sexual harassment of employees by other employees), and *Faragher v. City of Boca Raton*, 524 U.S. 775 (1998) (same).

11. *See* Rohde, *supra* note 6, at 216.

12. *See* Michael W. Giles, *Equivalent Versus Minimum Winning Opinion Coalition Size: A Test of Two Hypotheses*, 21 AM. J. POL. SCI. 405 (1977); R.W. Hoyer et al., *Some Problems in Validation of Mathematical and Stochastic Models of Political Phenomena: The Case of the Supreme Court*, 21 AM. J. POL. SCI. 381 (1977).

13. *See, e.g.*, WILLIAM H. RIKER & PETER C. ORDESHOOK, AN INTRODUCTION TO POSITIVE POLITICAL THEORY 177 (1973).

explain why there are so many unanimous decisions.

Finally, one might try to explain the data by looking at the value preferences of the Justices themselves. One might argue that there are many unanimous decisions because the Justices' values are clumped, so they tend to agree and reach consensus. Of course, one could make the converse argument as well: the pattern of alternating Republican and Democratic nominations means that Justices tend not to agree, resulting in many five-to-four decisions. Similarly, if we focused on what might be called "partner preferences" among Justices, we might conclude that certain Justices were more or less willing to join existing coalitions containing certain of their brethren.[14]

These substantive, value-based explanations suffer from several related problems. First, of course, every possible explanation may be contradicted by another possible explanation, and the conflict can be resolved only by looking at the data. Moreover, our data show that none of the value-based models are very plausible because value-based models can explain neither the consistency over four decades nor the prevalence of *both* unanimous and five-to-four opinions. Thus, our analysis suggests that research into individual preferences is not likely to yield useful results except at the margins.

II. THE MODEL

A. *Different Types of Models*

We begin by explaining what we mean by a model and what we hope to accomplish with it. Ours is not the typical social science model that hopes to *describe* the phenomenon in question. Rather, we present a model that is *functionally equivalent* to the phenomenon. By this we mean that the model will behave like the phenomenon—in this case, Supreme Court voting behavior—with respect to inputs and outputs, but without necessarily mimicking the actual phenomenon. For example, a model of the physics behind the moves of an expert

14. A spatial voting model of this sort is proposed in Lee Epstein & Carol Mershon, The Formation of Opinion Coalitions on the U.S. Supreme Court (1993) (unpublished manuscript, on file with authors). It is not clear how to use such a model to analyze the size of coalitions. Indeed, one of the few predictions that Epstein and Mershon make based on their model is that "[w]e see no reason for minimal winning coalitions to be especially frequent"—a prediction that is undermined by our data. *Id.* at 14.

Epstein and Mershon do not examine the question of coalition size very deeply and do not consider it a very interesting question: "In our view . . . the most important characteristic of coalitions to predict is not size but instead policy." *Id.* Our evidence would seem to indicate that the size of coalitions is more stable and more important than they believe.

billiards player does not purport to describe how the player actually analyzes the situation, but it does reproduce and predict the choices that the player makes. In other words, while the *player* does not typically calculate exact angles, trajectories, speed, and collision force, such calculations would yield the same results that the player achieves through instinct and experience.[15]

There is obviously a relationship between a descriptive model and a functionally equivalent one. A descriptive model must be functionally equivalent to be an accurate description—in other words, if a descriptive model does not adequately reflect the observed results, it fails as a description. A functionally equivalent model, on the other hand, need not be descriptive as long as it mirrors the observed results. Nevertheless, a functionally equivalent model is likely to have some descriptive relation to the actual phenomenon, if only because it is unlikely (although not impossible) that a model completely severed from the phenomenon would be functionally equivalent—at least over the long term. And a functionally equivalent model is also descriptive in the sense that it might serve as a sufficient explanation of the phenomenon, directing future research in particular directions and away from less fruitful explanations.

For example, a model that predicts the results of presidential elections by looking at the health of the economy is probably a fairly good functionally equivalent model. Although it is not descriptive—voters do not directly consult the latest government financial figures when casting their ballots—it bears some relation to how people evaluate incumbents. Moreover, such a model would suggest that future research focus on the details of the economy: What roles do inflation or unemployment play? Is it sufficient simply to look at the

---

15. This distinction between functionally equivalent and descriptive models is explored in MILTON FRIEDMAN, ESSAYS IN POSITIVE ECONOMICS 3–46 (1953). Although Friedman did not use the same terminology, he used the billiards example and defended what we are calling functionally equivalent models from attacks complaining of their descriptive flaws:

> It is frequently convenient to present such a hypothesis by stating that the phenomena it is desired to predict behave in the world of observation *as if* they occurred in a hypothetical and highly simplified world containing only the forces that the hypothesis asserts to be important. . . . Such a theory cannot be tested by comparing its "assumptions" directly with "reality." Indeed, there is no meaningful way in which this can be done. Complete "realism" is clearly unattainable, and the question whether a theory is realistic "enough" can be settled only by seeing whether it yields predictions that are good enough for the purpose in hand or that are better than predictions from alternative theories.

Id. at 40–41. For a more contemporary discussion of these two different views of modeling in the context of rational choice theory, see DONALD P. GREEN & IAN SHAPIRO, PATHOLOGIES OF RATIONAL CHOICE 30 (1994).

stock market? Are there ways in which people's *perceptions* of the health of the economy can be influenced, with a consequent effect on the election? On the other hand, a model that predicts the results of presidential elections by means of astrology might be successful as a functional equivalent for a few elections, but it is not likely to work for very long, precisely because it bears *no* relation to the phenomenon itself.[16]

Despite some relationship between a descriptive model and a functionally equivalent one, however, they are not identical, and thus should be judged by different criteria. A descriptive model should give an accurate and thorough account of the phenomenon, including all of the major factors affecting the phenomenon. By this standard, our model is quite poor. The standard is irrelevant, however, because our model is not intended to be descriptive.

A functionally equivalent model, by contrast, need not provide any substantive description. The value of the model is in the numbers. For a functionally equivalent model to be valid, the numbers must be consistent with those produced by the phenomenon in the real world and must fit with all other available data. For a functionally equivalent model to be useful and interesting, however, it must do more: it should reveal relationships that were not previously apparent and that point to nonobvious, empirically testable relationships in the phenomenon itself. By this standard, our model is quite successful.

Our model posits that the size of a final coalition is governed by a probabilistic process in which the transition probability—that is, the likelihood that at least one Justice will join the existing coalition—is dependent only on the size of the coalition. Based on this assumption, we construct a Markov process with absorbing states, with which we model the coalition formation of the Court. Calculating the transition probabilities shows that in general they *increase* with the size of the majority coalition.[17] This increase suggests that, consistent with some of the prior research,[18] Justices are increasingly less likely to remain in smaller minorities.

The idea behind our Markov model is that each Justice not in a currently existing majority must decide whether to join that majority.

---

16. A valid descriptive model for presidential elections would be extremely complex, taking into account the myriad factors that voters actually consider in casting their ballots. No one has so far been able to produce such a model, although presidential candidates continually seek one.

17. *See infra* Part III.

18. *See* Dorff & Brenner, *supra* note 2, at 768.

We assume that each Justice makes her decision independently and that the decision does not depend on which particular Justices are in or outside of the majority. As we will see, the data generally support these assumptions, although some minor variations may be due to the personalities of individual Justices and their interactions with one another. Indeed, our data suggest that it might be useful and interesting to examine individual Justices and their interactions in order to explain the minor fluctuations between Terms. Nevertheless, the numbers would not work out the way they do if Justices were influenced in an important way by the presence or absence of other Justices in the coalition.

The assumption that the formation of a coalition is a process rather than an event—and that it is influenced by knowledge of the size of the coalition existing at any given time—is supported by how the Supreme Court operates in practice. The Justices certainly have numerous opportunities to assess the size of the majority coalition and to tailor their own votes accordingly. In addition to informal discussions among chambers, all of the Justices present their views on a case at the conference before a final vote is taken;[19] moreover, Justices not infrequently change their votes even after the conference.[20]

Note that we do not suggest that the individual Justices actually *perform* any mathematical calculations or even consciously rely on discomfort with being in a small minority. We contend only that this model explains the data better than any other model and, as we discuss in Part III, that the model reveals a surprising result that is not apparent from using other models.

At this point, however, a reader might ask why our model is important if it does not rely on what the Justices actually think about; another way to put it is to ask why a functionally equivalent model is useful if it is not also descriptive. Such a model is useful because it rules out certain descriptive explanations and makes others more plausible. For example, we know from our data that (except possibly in the case of a transition from a five-Justice to a six-Justice majority) the dissenters, not the majority, drive the increase in size. We also know that any explanation that relies primarily on individual personalities and preferences is unnecessarily complicated, although such individual differences might serve to add interesting insights to our model.

---

19. *See, e.g.*, EPSTEIN & KNIGHT, *supra* note 6, at 65–79.

20. *Cf. id.* at 9 (providing an example of judicial voting in a particular case).

B. The Details of the Model

The model works in the following way. Suppose we currently have a majority coalition of size k, where $5 \leq k \leq 8$.[21] Assume that each of the $9 - k$ Justices not in the coalition will join the big coalition with a probability p_k, a probability that is dependent only on the size of the current majority coalition. We call this probability the *joining probability*. The non-coalition (or minority) Justices will make their determinations simultaneously and independently. If none of the Justices join the coalition, the process is terminated with a majority coalition of size k. If all of the Justices join the coalition, the process is terminated with a unanimous decision. If some but not all of the Justices choose to join the coalition, the process starts again with a new majority coalition of the new size.

As an example, suppose that we begin with 6 Justices currently in the majority and that $p_6 = 1/3$. This means that for each of the 3 Justices not in the majority, the probability that he will join the majority is 1/3. The probability that none of them will join, thus yielding a final coalition of 6, is $(1 - 1/3)^3 = 8/27$. The probability that all of them will join, thus yielding a unanimous opinion, is $(1/3)^3 = 1/27$. Finally, the probability that the coalition will have grown, but not to unanimity, meaning that the process will begin again with a new coalition size, is $1 - 8/27 - 1/27 = 2/3$.

If we have all of the joining probabilities for any given Term, we can calculate the exact distribution of coalition sizes. We assume that all coalitions start at five. From p_5 we can calculate the percentage of five-Justice coalitions that will, in the first round of the "game," stay as five-Justice coalitions[22] or become six-Justice coalitions,[23] seven-Justice coalitions,[24] eight-Justice coalitions,[25] or nine-Justice coalitions.[26] Then we take the six-Justice coalitions and use p_6 to calculate what percentage will turn into seven-, eight-, and nine-Justice coalitions in the second round and repeat the process for p_7 and p_8. At this point, we simply tally up our accumulated percentages from each round for each coalition size. Figure 1 depicts the process in graphic form.

21. If $k = 9$, we have a unanimous decision, and the process is terminated. If $k < 5$, we did not start with a majority.

22. $(1 - p_5)^4$.

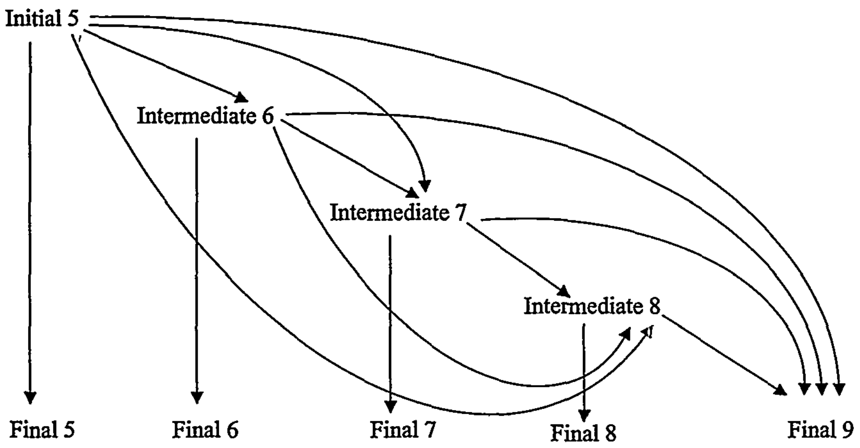
23. $4 \times p_5 \times (1 - p_5)^3$.

24. $6 \times p_5^2 \times (1 - p_5)^2$.

25. $4 \times p_5^3 \times (1 - p_5)$.

26. p_5^4.

FIGURE 1
THE COALITION FORMING PROCESS



(Intermediate means that, although the coalition size has been reached, the game is not over.)

Notice that any given set of joining probabilities will yield only one distribution pattern. Note also that neither the joining probabilities nor their relationships to each other are obvious from the distributions; they must be calculated. Thus, the joining probabilities can reveal more information about the Justices' interactions than does a simple table of coalition-size distributions. The joining probabilities also quantify and specify with better exactitude the tendencies that have been noted by the scholars who have investigated the paper record and have concluded that vote-changes become more likely as the size of the minority decreases.[27]

It turns out that this particular Markov process can also be solved in reverse.[28] That is, from a given distribution, a unique set of joining probabilities can be calculated. In Part III, we calculate the joining probabilities for twelve Supreme Court Terms. Although it is theoretically possible that constructing the joining probabilities in this

27. *See* Dorff & Brenner, *supra* note 2, at 768.

28. The details of the computation are given in Appendix A. Not all Markov processes can be solved in reverse; indeed, most of them cannot. In a general Markov process, the number of different transition probabilities exceeds the number of absorbing states, so it is mathematically impossible to solve for the probabilities in terms of the final distribution of outcomes.

way might yield "joining probabilities" that are negative (which would indicate a serious flaw in the model), the actual data never produced such probabilities. This result confirms that the data are consistent with the model. If the data had produced negative "joining probabilities"—as in theory they might have—then we would have had to abandon the model.

C. *Methodology*

Before we turn to the results of our computation, we must discuss several methodological issues. In applying the model, we assume that the process starts with a coalition size of five. Why not with a larger or smaller coalition? First, starting with a minimal winning coalition is in the spirit of other cooperative game theory analyses. For instance, the Shapley-Shubik model for measuring voting strength assumes that the votes are cast sequentially, and the voter who casts the winning vote is seen as the pivotal voter.[29] While perhaps not realistic, Shapley and Shubik's assumption of sequential votes represents a conventionally accepted way to analyze the strength of various voters.[30] In the same way, we assume that coalitions are formed sequentially, starting from the minimal winning coalition of five votes. Note also that our model has nothing to say on the question of how this original majority coalition comes into being.

There is a second, more practical, reason to begin with a coalition of five rather than a larger or smaller coalition. How would one identify the "original" coalition size in any given case? If we chose the size of the majority coalition in the conference vote, that would ignore all information a Justice acquired *before* voting at conference—including not only informal conversations prior to the conference, but the conference discussion itself. It is not implausible that a Justice waiting her turn to vote might see a majority coalition in the process of forming and take that information into account. Thus, the conference vote might itself depend on joining probabilities.

Perhaps the most problematic aspect of beginning with a five-Justice coalition is that some interesting cases involve larger initial majorities that suffer defections; indeed, one will occasionally read a majority opinion that obviously began life as a dissent, or vice versa.

---

29. *See* L.S. Shapley & Martin Shubik, *A Method for Evaluating the Distribution of Power in a Committee System*, 48 AM. POL. SCI. REV. 787, 788 (1954); *see also* Paul H. Edelman & Jim Chen, *The Most Dangerous Justice: The Supreme Court at the Bar of Mathematics*, 70 S. CAL. L. REV. 63, 74–75 (1996) (noting that not all fifth votes are pivotal because some threats to dissent are not credible).

30. *See, e.g.*, ALAN D. TAYLOR, MATHEMATICS AND POLITICS 65–71 (1995).

Our model says nothing about the probability that a Justice might leave a majority coalition. It would not be difficult to calculate the probability that, say, a seven-Justice coalition will become a six-Justice coalition, but doing so seems counterintuitive because it would require starting with an assumption that decisions begin as unanimous and then deteriorate into smaller majorities. Thus, the decision not to focus on defections was partly due to the implausibility of this assumption, but also to the fact that defections from majority coalitions are likely to be less frequent than decisions to join them. For one thing, a credible threat to desert[31] might influence a majority opinion author to make changes. Moreover, our data show an increasing unwillingness to be in a small minority, suggesting that the most likely defection is from a five-Justice majority—which simply changes our starting point from one five-Justice majority to a different one.

A second methodological question asks how we measure the size of the majority coalition. Do we look at the size of the coalition producing the *decision* or the size of the coalition joining the majority *opinion*? We believe that the size of the opinion coalition is more meaningful because Justices who concur only in the result can disagree substantially with the majority. Thus, focusing on the number of Justices concurring only in the result can mask tremendous differences in reasoning and conclusions.[32]

As an example, consider the case of *44 Liquormart, Inc. v. Rhode Island*.[33] Although there were no dissenters from the judgment of the Court invalidating Rhode Island's restrictions on liquor advertising, there was no majority opinion, and the Justices disagreed about both the breadth and the vitality of the relevant precedent.[34] To consider the case as a unanimous decision would grossly misrepresent the result. More to the point for our analysis, no Justice was ever presented with a majority opinion as a *fait accompli*, and thus our "game" would never have started. Just to confirm the validity of our results, however, we recalculated the 1995 and 1996 Terms, measuring the size of the majority by the size of the decision. As expected, the joining probabilities went up, but not by very much, and the overall

---

31. By credible, we mean that it would be plausible for the Justice in question to join the minority coalition. *See* Edelman & Chen, *supra* note 29, at 83.

32. *See id.* at 79.

33. 517 U.S. 484 (1996).

34. *See id.* at 488–516; *id.* at 517–18 (Scalia, J., concurring in part and concurring in the judgment); *id.* at 518–28 (Thomas, J., concurring in Parts I, II, VI, and VII, and concurring in the judgment); *id.* at 528–34 (O'Connor, J., concurring in the judgment).

pattern remained the same.[35]

A third methodological issue concerns the selection of the twelve particular Supreme Court Terms. In each decade from the 1960s to the 1990s, we selected three consecutive Terms of the same Court—that is, three consecutive Terms without a change in personnel.[36] A range of three Terms gave us a deeper picture of each Court, and including in our study four decades of decisions ensured that the results did not depend on the idiosyncrasies of any particular collection of Justices. When there were more than three consecutive Terms of the same Court, we chose the last three in order to reflect a more mature Court with more settled interactions. We therefore included the following Terms: 1962–1964,[37] 1972–1974,[38] 1983–1985,[39] and 1995–1997.[40]

The last methodological point is to describe in detail how we tallied the cases.[41] For the reasons suggested above, cases without a

---

35. We give the joining probabilities from Table 1, followed in parentheses by the joining probabilities that would result if we used the size of the majority *decision* rather than the majority *opinion*:

1995: $p_5 = .37\ (.39), p_6 = .30\ (.36), p_7 = .44\ (.40), p_8 = .75\ (.75)$

1996: $p_5 = .32\ (.33), p_6 = .27\ (.31), p_7 = .43\ (.55), p_8 = .67\ (.77)$

We chose these two Terms for convenience. In beginning our research, we had started with these two Terms and had initially included enough information in the database to make recalculation easy. For the other Terms, we would have had to go back and recount the cases.

36. We were not able to collect data from the 1950s because there were not three consecutive years of a stable Court during that decade.

37. During this period, the Court consisted of Chief Justice Earl Warren and Associate Justices Hugo L. Black, William O. Douglas, Tom C. Clark, John M. Harlan, William J. Brennan, Jr., Potter Stewart, Byron R. White, and Arthur J. Goldberg. *See* 371 U.S. III; *see also* GERALD GUNTHER & KATHLEEN SULLIVAN, CONSTITUTIONAL LAW, at B-5 to B-6 (13th ed. 1997) (chart).

38. During this period, the Court consisted of Chief Justice Warren E. Burger and Associate Justices William O. Douglas, William J. Brennan, Jr., Potter Stewart, Byron R. White, Thurgood Marshall, Harry A. Blackmun, Lewis F. Powell, Jr., and William H. Rehnquist. *See* 409 U.S. III; *see also* GUNTHER & SULLIVAN, *supra* note 37, at B-5 to B-6 (chart).

39. During this period, the Court consisted of Chief Justice Warren E. Burger and Associate Justices William J. Brennan, Jr., Byron R. White, Thurgood Marshall, Harry A. Blackmun, Lewis F. Powell, Jr., William H. Rehnquist, John Paul Stevens, and Sandra Day O'Connor. *See* 464 U.S. III; *see also* GUNTHER & SULLIVAN, *supra* note 37, at B-6 to B-7 (chart).

40. During this period, the Court consisted of Chief Justice William H. Rehnquist, and Associate Justices John Paul Stevens, Sandra Day O'Connor, Antonin Scalia, Anthony M. Kennedy, David H. Souter, Clarence Thomas, Ruth Bader Ginsburg, and Stephen Breyer. *See* 516 U.S. III; *see also* GUNTHER & SULLIVAN, *supra* note 37, at B-6 to B-7 (chart).

41. All counts were done by hand by the authors. We looked at the listing of Justices provided at the beginning of each opinion in *U.S. Law Week* (for the 1995–1997 Terms) or

majority opinion were excluded from the data. Because we were trying to determine whether the size of the majority or the size of the minority was more significant, any case in which fewer than nine Justices participated was also excluded. We could not know, for example, whether a seven-to-one decision was mathematically equivalent to a seven-to-two decision or to an eight-to-one decision.[42]

We also excluded all cases in which the majority opinion was per curiam.[43] Certainly, many per curiam decisions should be excluded: it would not have been appropriate, for example, to include per curiam decisions dismissing a writ of certiorari as improvidently granted. Other per curiam decisions probably should have been included in the count, as they addressed substantive issues. Nevertheless, the judgment about whether a per curiam case was important enough to count could not be made consistently—and in any case, excluding them did not make much difference.[44]

---

the *United States Supreme Court Reports: Lawyer's Edition (2d)* (for the earlier Terms). Any Justice listed as joining the majority opinion was counted as part of the coalition, regardless of whether he or she also wrote a concurrence. Any ambiguities were resolved by looking at the individual opinions. All legal judgment calls were made by Professor Sherry, who is solely responsible for them.

42. Excluding cases decided by smaller Courts may explain the peculiar results from the 1964 Term, which had an exceptionally high number of missing Justices.

43. We made one exception to this rule. In *Bazemore v. Friday*, 478 U.S. 385 (1986), the per curiam opinion simply announced the two parts of the judgment and noted that the first part was reached for the reasons given in Justice Brennan's unanimous concurring opinion and the second for the reasons given in Justice White's five-Justice concurring opinion. We made a judgment call, concluding that the per curiam format was used for convenience only, and counted it as two separate majorities of sizes nine and five, respectively.

44. Although for the later years of our study we might have relied on the judgments made in the annual Supreme Court issue of the *Harvard Law Review*, that source was not adequate for the earlier Terms. To confirm the validity of our results, however, we recalculated the 1995 and 1996 Terms including the per curiam cases identified as substantial by the Harvard survey and found that the results were virtually identical. Thus, for those two years, we can compare the results in Tables 1 and 2, which exclude all per curiam decisions, with the results we initially obtained, which include those per curiam decisions counted by the Harvard survey. There is virtually no change. Here are the numbers for the 1995 and 1996 Terms (the figures in parentheses are those obtained by using the Harvard count):

1995: $p_5 = .37$ (.36), $p_6 = 30$ (.30), $p_7 = .44$ (.41), $p_8 = .75$ (.76)
1996: $p_5 = .32$ (.33), $p_6 = .27$ (.25), $p_7 = .43$ (.43), $p_8 = .67$ (.71)

The Harvard survey added 4 per curiam cases for the 1995 Term, *see Leading Cases*, 110 HARV. L. REV. 135, 367 n.a (1996), and 6 per curiam cases for the 1996 Term, *see Leading Cases*, 111 HARV. L. REV. 197, 431 n.a (1997). We chose these two Terms for convenience. In beginning the research, we had started with these two Terms and had initially relied on the Harvard survey until we discovered that it was not adequate for the earlier years. Thus our initial data for the two Terms *did* include per curiam decisions, so when we recalculated without the per curiam opinions (for the data that appear in Tables 1 and 2), we had a ready comparison.

Finally, some cases produced multiple majorities of different sizes, with various Justices joining some parts of the majority opinion but not others. For most of these cases, we tallied the case as representing a single instance of each size that occurred. In *Koon v. United States*,[45] for example, the question was whether the court of appeals correctly overturned a district judge's downward departure from the United States Sentencing Guidelines.[46] The Supreme Court decided unanimously that the court of appeals applied an incorrect standard of review.[47] In applying the correct standard, however, the Court itself fractured: of the four factors cited by the district court to support the downward departure, the Court rejected one of the factors unanimously,[48] rejected a second factor by a vote of eight to one,[49] and accepted two factors by a vote of six to three.[50] We thus counted *Koon* three times: once as a unanimous opinion, once as an eight-Justice majority, and once as a six-Justice majority. We only counted cases as representing multiple majorities, however, if the portion of the opinion that a Justice refused to join was legally significant.[51] For example, in *Bank One Chicago v. Midwest Bank & Trust Co.*,[52] Justice Scalia joined almost all of the majority's opinion interpreting the Expedited Funds Availability Act,[53] but refused to join the portion of the opinion that looked at legislative history to confirm the majority's interpretation.[54] We did not count that dispute as yielding a separate coalition size.

45. 518 U.S. 81 (1996).

46. *See id.* at 85.

47. *See id.* at 91.

48. *See id.* at 111.

49. *See id.* at 109–11; *id.* at 114 (Stevens, J., concurring in part and dissenting in part).

50. *See id.* at 111–12; *id.* at 114–18 (Souter, J., concurring in part and dissenting in part); *id.* at 118–19 (Breyer, J., concurring in part and dissenting in part).

51. Again, there was one unusual exception. In *Strickland v. Washington*, 466 U.S. 668 (1984), Justice Brennan joined the majority's opinion but not its judgment. *See id.* at 701 (Brennan, J., concurring in part and dissenting in part). He agreed with the majority's substantive analysis and conclusion (that a criminal defendant had received constitutionally adequate counsel), but dissented from the judgment on the ground—not raised by the parties or considered by any other Justice—that the death penalty violated the Eighth Amendment. *See id.* (Brennan, J., concurring in part and dissenting in part). Although technically a legally significant difference (after all, he, unlike the majority whose opinion he joined, would have reversed the court below), the dispute was not really part of the case, and so we included him in the majority.

52. 516 U.S. 264 (1996).

53. 12 U.S.C. §§ 4001–4010 (1994).

54. *See Bank One Chicago*, 516 U.S. at 279–80 (Scalia, J., concurring in part and concurring in the judgment).

III. THE RESULTS

A. *Joining Probabilities*

Although the distribution of majority coalition sizes varied from year to year, as did the particular joining probabilities, there was one constant. With the exception of slight variations in the 1964 and 1973 Terms, $p_8 > p_7 > p_5 > p_6$. In other words, the probability that a minority Justice would join an eight-Justice majority was the largest, followed by the probability of joining a seven-Justice majority, followed by the probability of joining a *five*-Justice majority. Surprisingly, in each of the twelve Terms the most stable coalition size is not five but six. By the most stable, we mean that minority Justices are least likely to join a six-Justice coalition.

The joining probabilities for each Term are given in Table 1, and the aggregate joining probabilities for each of the four Courts over three Terms are given in Table 2.[55]

TABLE 1
JOINING PROBABILITIES BY TERM

Term	p_5	p_6	p_7	p_8
1962	.34	.20	.40	.42
1963	.32	.23	.45	.56
1964	.39	.17	.27	.33
1972	.27	.16	.38	.45
1973	.26	.15	.54	.40
1974	.29	.13	.45	.55
1983	.30	.21	.55	.59
1984	.30	.19	.35	.67
1985	.23	.20	.45	.55
1995	.37	.30	.44	.75
1996	.32	.27	.43	.67
1997	.33	.19	.56	.76

55. We provide the raw data from which these tables were constructed—that is, the distributions of coalition sizes—in Appendix B.

TABLE 2
JOINING PROBABILITIES BY COURT[56]

Court	p_5	p_6	p_7	p_8
1962–1964	.34	.20	.37	.45
1972–1974	.27	.15	.45	.47
1983–1985	.27	.20	.46	.60
1995–1997	.34	.25	.48	.73

What is especially interesting is that the pattern of joining probabilities remains consistent from Court to Court. Indeed, as Table 2 shows, if we aggregate the joining probabilities for each Court, the probabilities themselves are remarkably similar from decade to decade and from Court to Court. This is true despite the large differences in the actual distributions. For example, the percentage of unanimous opinions ranges from a low of 18 (in 1964 and 1972) to a high of 46 (in 1995); the percentage of five-to-four opinions ranges from 14 (in 1964) to 36 (in 1985).[57]

B. Comments

Given the persistence of the pattern, we can comfortably use it to draw some conclusions. First, the relatively high values of p_8 and p_7 help explain some previous research results and reject some theories. Both theoretical analysis of the marginal utility of each additional vote and empirical research on the Justices' behavior suggest that the willingness of a majority to compromise for additional votes should decline with the increasing size of the majority.[58] Even if, as some scholars have argued, majorities have an incentive to bargain because supermajorities have enhanced influence, that incentive should decline with the increasing size of the majority.[59] Thus, if the desires

56. These aggregated probabilities were calculated by counting the total numbers of each coalition size during the entire three-year period and applying the model to the aggregated data. Thus, the aggregated probability is *not* simply an average of the probabilities of individual years.

57. The distributions and percentages are given in Appendix B.

58. *See, e.g.*, WALTER F. MURPHY, ELEMENTS OF JUDICIAL STRATEGY 65 (1964); Wahlbeck et al., *supra* note 1, at 312.

59. *See, e.g.*, Frank B. Cross, *The Justices of Strategy*, 48 DUKE L.J. 511, 556 (1999) (making this argument and conceding that the power to command changes in the majority opinion declines with each additional vote).

of the existing majority were the significant factor in attracting more Justices to the coalition, we would expect the joining probabilities to decline as the size of the majority increases.

Our data, however, show exactly the opposite. The high values of p_8 and p_7 suggest that, even if the majority is unwilling to compromise, the last few remaining Justices not in the coalition are very interested in joining. And being in a minority of one is even less desirable than being in a minority of two. Our data thus demonstrate what we call a "bandwagon effect." The bandwagon effect is consistent with earlier work showing that lone dissenters are more likely than any other Justices to switch their votes.[60] Our analysis therefore indicates that the focus on the growth of coalitions generally should be on the size of the *minority* and its relative desirability, rather than on the size of the majority and the power that it wields or on the value to the Court of supermajorities.

Note that while our data document and measure the bandwagon effect, the data do not explain its causes. Scholars have suggested several possibilities. One is that, lacking much power to influence the development of doctrine, minority Justices are unwilling to expend the time and effort necessary to write a separate opinion.[61] Another is that at least some Justices "like to win (or to be perceived as 'winners'),"[62] and a third is simple collegiality.[63] Whatever the reason, our data confirm that Justices in small minorities are drawn toward joining the majority opinion.

The model thus suggests that the distribution of majority opinion coalition sizes can be explained solely as a consequence of the fact that the joining probabilities increase with the size of the majority. That is, earlier empirical research indicating an increasing willingness to join large majorities[64] is by itself sufficient to explain the actual distribution of opinion sizes. Future empirical research now might be directed toward the motivations of the dissenting Justices and the

60. *See, e.g.*, Brenner et al., *supra* note 2, at 249–52; Dorff & Brenner, *supra* note 2, at 773; Maltzman & Wahlbeck, *supra* note 2, at 587–89. Our data do conflict with one finding of Brenner et al.: they concluded that Justices in a four-Justice minority were least likely to switch votes, and our data show that a *three*-Justice minority is in fact the most stable. *See infra* text accompanying notes 66–69. Further research is warranted here.

61. *See* Frank B. Cross, *Political Science and the New Legal Realism: A Case of Unfortunate Interdisciplinary Ignorance*, 92 NW. U. L. REV. 251, 305–06 (1997).

62. Tracey E. George, *Developing a Positive Theory of Decisionmaking on U.S. Courts of Appeals*, 58 OHIO ST. L.J. 1635, 1661 (1998).

63. *See* Lewis A. Kornhauser & Lawrence G. Sager, *The One and the Many: Adjudication in Collegial Courts*, 81 CAL. L. REV. 1, 8 (1993).

64. *See supra* note 2 (citing sources).

causes of the bandwagon effect. It might also focus on whether particular Justices are more or less susceptible to the bandwagon effect.

The bandwagon effect has increased noticeably over the years, at least with regard to lone dissenters. Between the 1960s Court and the 1990s Court, p_8 increased by more than fifty percent (from .45 to .73), while the other joining probabilities remained constant or increased only slightly. This finding may be due to the personalities or work habits of individual Justices or to a strengthening institutional norm of collegiality.[65] Another possibility is the change in institutional norms regarding individual expression and the resulting workload. As Justices have come to rely more heavily on such devices as the "cert. pool,"[66] their comfort level with collaborative efforts may have increased. Finally, the increase in p_8 might be due to a change in opinion writing patterns. It used to be relatively common for Justices to dissent or to concur only in the judgment *without* writing a separate opinion. Now Justices rarely do so, obviously increasing the workload of those who decline to join the majority opinion. Further research might be warranted to examine the relationship between the value of p_8 and the number of dissents and concurrences in judgment without opinion.

A second, and perhaps more interesting, aspect of the data is the relationship between p_6 and p_5. Both the bandwagon effect we document here and standard game theory predictions of minimal winning coalitions suggest that p_5 should be lower than p_6, but the data consistently show the opposite. Indeed, a non-majority Justice is only about two-thirds as likely to join a six-Justice coalition as to join a five-Justice coalition—a substantial difference in an unexpected direction. What might account for this anomaly?

One possibility is that, while the size of the minority is the significant factor in most instances, a five-to-four majority has particular incentives to bargain in order to attract at least one additional vote. Perhaps this slenderest of majorities perceives itself as particularly vulnerable to overruling by Congress in statutory cases

65. For more on institutional norms on the Supreme Court, see EPSTEIN & KNIGHT, *supra* note 6, at 118.

66. The cert. pool involves letting a group of law clerks screen petitions for certiorari and circulate memoranda to the chambers of *all* the Justices who are members of the pool. Begun during the Burger Court era by five Justices, the cert. pool now includes all the Justices except Justice Stevens. *See* David M. O'Brien, *Join-3 Votes, the Rule of Four, the Cert. Pool, and the Supreme Court's Shrinking Plenary Docket*, 13 J.L. & POL. 779, 790, 799–800 (1997).

or by subsequent Courts in constitutional cases.[67] A five-Justice majority also might be concerned about the possibility of defections, which obviously have a more devastating effect on a minimally winning majority. In any case, it is clear that there is *something* unique about a five-Justice coalition that produces a higher than expected joining probability. Again, further empirical research is needed, this time utilizing the various Justices' papers.[68] While prior research has focused, incorrectly, as our data show, on whether unanimity is more likely in important cases,[69] it might now be useful to determine whether important cases are more likely to be six-to-three decisions than five-to-four decisions.

CONCLUSION

Over the course of four decades, the Supreme Court changes in many ways. The personalities of the individual Justices and their interactions with one another are different for each Court. Some Courts are more collegial, some Justices more iconoclastic. Each Term presents a different mix of difficult or controversial cases. Nevertheless, despite these differences, our mathematical model shows several constants. First, when Justices join a majority coalition, it is more likely to be because of their own desire to join than because of any accommodations made by the majority. Second, it appears that over time, the Justices have increasingly viewed being a lone dissenter as quite undesirable. Finally, the most interesting result of applying our model to the data is that, despite minority Justices' marked preference for being in a larger rather than a smaller minority, a majority coalition is more likely to attract a sixth Justice than it is to attract a seventh. This result suggests that a majority of five is sufficiently willing to negotiate for additional votes to overcome the minority Justices' disinclination to abandon what would

---

67. For other suggestions along this line, see William N. Eskridge, Jr., *Overriding Supreme Court Statutory Interpretation Decisions*, 101 YALE L.J. 331 (1991). There is also the somewhat remote possibility that lower courts will be more hostile toward five-to-four decisions. *See, e.g.*, Paper Converting Mach. Co. v. Magna-Graphics Corp., 745 F.2d 11, 17 (Fed. Cir. 1984) ("We must be cautious in extending five to four decisions by analogy . . . ." (citations omitted)); United States v. Kennesaw Mountain Battlefield Ass'n, 99 F.2d 830, 833–34 (5th Cir. 1938) ("We agree that it is not controlling for the further reason, that that case, decided as it was by a closely divided court, is authority only for its own facts, and those facts are not present here.").

68. Epstein and Knight have taken this approach and showed that bargaining does occur. *See* EPSTEIN & KNIGHT, *supra* note 6, at 56–111. It remains to examine whether a five-Justice majority is indeed more eager to bargain and, if so, why.

69. *See supra* note 6 (citing sources).

otherwise be a comfortably large minority.

Our model thus presents a clearer picture of Supreme Court coalition formation and answers questions left open by previous research. It confirms that, with the exception of the transition from a five-to-four decision to a six-to-three decision, future research should focus on the inclinations of the minority Justices rather than on the motivations of the majority Justices. On the other hand, it offers potential statistical support for those who explore the peculiar vulnerabilities—and consequent negotiation tactics—of five-Justice majorities. Our research also suggests fruitful avenues for further research, including an investigation of individual Justices' relative disinclinations to remain in small minorities and an examination of the relationship between opinion writing norms and the willingness to remain in a small minority.

APPENDIX A: THE MODEL

In this Appendix, we describe the model and explain the mathematics involved in its solution. The model is a basic Markov process with absorbing states. For some background, the reader might consult William Feller, *An Introduction to Probability Theory and Its Applications*, ch. 15 (3d ed. 1968).

The states of our model will be denoted by the set

$$S = \{5, 6, 7, 8\} \cup \{\bar{5}, \bar{6}, \bar{7}, \bar{8}, \bar{9}\}$$

where the states $\bar{i}$ will be absorbing states. Let p_5, p_6, p_7, p_8 be probabilities where p_i is the probability that a Justice will join a coalition of i other Justices. If we assume that Justices not in the coalition make decisions to join independently, then the probability of exactly i of the $9 - j$ Justices joining a coalition of j Justices is

$$B(i, j) = \binom{9-j}{i} p_j^i (1 - p_j)^{9-j-i}$$

One can now define the non-zero transition probabilities. Let $P(\alpha, \beta)$ be the probability of going to state β from state α. Then

$$P(i, j) = B(j - i, i) \text{ for } 5 \leq i \leq j \leq 8 \quad (1)$$
$$P(i, \bar{i}) = B(0, i) \text{ for } 5 \leq i \leq 8 \quad (2)$$
$$P(i, \bar{9}) = B(9 - i, i) \text{ for } 5 \leq i \leq 8 \quad (3)$$
$$P(\bar{i}, \bar{i}) = 1 \text{ for } 5 \leq i \leq 9 \quad (4)$$

and the remaining transition probabilities are 0.

It is now a relatively simple matter to solve for the probability of starting in state 5 (that is, with a coalition of 5 Justices) and ending, eventually, in a state $\bar{i}$ for $i = 5, 6, \ldots, 9$ (that is, the final coalition has size i). Call this probability C_i. The results are:

$C_5 = B(0,5)$

$C_6 = B(0,6)\, B(1,5)$

$C_7 = B(0,7)\, [B(2,5) + B(1,5)\, B(1,6)]$

$C_8 = B(0,8)\, [B(3,5) + B(2,5)\, B(1,7) + B(1,5)\, [B(2,6) + B(1,6) B(1,7)]]$

$C_9 = B(4,5) + B(3,5)\, B(1,8) + B(2,5)\, [B(2,7) + B(1,7)\, B(1,8)] + B(1,5)\, [B(3,6) + B(2,6)\, B(1,8) + B(1,6)\, [B(2,7) + B(1,7)\, B(1,8)]]$

The reader will observe that the probability C_i is a function only of the probabilities p_j where $5 \leq j \leq i$. Thus, C_5 is solely a function of p_5, C_6 is solely a function of p_5 and p_6, and so on. Hence, the values of the C_i's completely determine the probabilities $p_5, \ldots, p_8$. Those equations will not be explicitly written down here.

We should note that not all probabilities C_i are compatible with this model. For example, if $C_6 \div B(1,5)$ is larger than one, then solving for p_6 will result in a negative number, clearly not the probability of anything. The fact that the data we have collected is always compatible and that the solutions for the joining probabilities are legitimate probability values lends credibility to the model.

APPENDIX B
DISTRIBUTIONS OF COALITION SIZES

YEAR	MAJORITY COALITION	NUMBER	PERCENT (ROUNDED)
1962	5	19	19
	6	20	20
	7	16	16
	8	20	20
	9	23	23
1963	5	22	22
	6	19	19
	7	14	14
	8	16	16
	9	31	30
1964	5	11	14
	6	16	20
	7	20	25
	8	19	24
	9	14	18
1962–1964	5	52	19
	6	55	20
	7	50	18
	8	55	20
	9	68	24

YEAR	MAJORITY COALITION	NUMBER	PERCENT (ROUNDED)
1972	5	39	28
	6	34	25
	7	20	14
	8	20	14
	9	25	18
1973	5	39	30
	6	34	26
	7	10	8
	8	20	15
	9	28	21
1974	5	26	25
	6	29	28
	7	12	11
	8	13	12
	9	25	24
1972–1974	5	104	28
	6	97	26
	7	42	11
	8	53	14
	9	78	21

YEAR	MAJORITY COALITION	NUMBER	PERCENT (ROUNDED)
1983	5	36	24
	6	31	21
	7	13	9
	8	20	13
	9	50	33
1984	5	22	24
	6	20	22
	7	16	18
	8	9	10
	9	24	26
1985	5	53	36
	6	32	22
	7	15	10
	8	16	11
	9	31	21
1983–1985	5	111	29
	6	83	21
	7	44	11
	8	45	12
	9	105	27

YEAR	MAJORITY COALITION	NUMBER	PERCENT (ROUNDED)
1995	5	11	15
	6	9	13
	7	11	15
	8	8	11
	9	33	46
1996	5	19	22
	6	14	16
	7	13	15
	8	11	13
	9	31	35
1997	5	18	20
	6	19	21
	7	8	9
	8	8	9
	9	39	42
1995–1997	5	48	19
	6	42	17
	7	32	13
	8	27	11
	9	103	41