Inside the U.S. Supreme Court: The Reliability of the Justices' Conference Records

Forrest Maltzman; Paul J. Wahlbeck

The Journal of Politics, Vol. 58, No. 2 (May, 1996), 528-539.

Stable URL:
http://links.jstor.org/sici?sici=0022-3816%28199605%2958%3A2%3C528%3AITUSCT%3E2.0.CO%3B2-7

The Journal of Politics is currently published by Southern Political Science Association.

Inside the U.S. Supreme Court: The Reliability of the Justices' Conference Records

Forrest Maltzman and Paul J. Wahlbeck
George Washington University

Judicial scholars are making increased use of data from the justices' personal papers. In the face of comments by justices questioning the reliability of this information and, perhaps, skepticism among judicial scholars about the reliability of some justices' records, it is important to explore the reliability of data drawn from their papers. To do so, we examine the conference records for the 1967 and 1968 terms of the Supreme Court. For these years, the papers of four justices, Earl Warren, William Douglas, William Brennan, and Thurgood Marshall, are available at the Library of Congress. We found that their records are substantially accurate and reliable. A logit model of conference recording error finds that mistakes are not simply random noise, but can be explained as a function of case complexity, the length of joint service with the voting justices, tentative conference positions, and Court consensus.

Most of the Supreme Court's decision-making process takes place behind closed doors. Twice a week, the justices meet in private conference to decide which cases they will hear and to discuss the merits of the cases that were argued earlier that week. The only people allowed access to these conferences are the justices themselves (Rehnquist 1987). No clerks or members of the support staff are present. Consequently, the public and the scholarly community do not have access to the justices' internal deliberations. We are instead limited to a relatively narrow public view of the process: certiorari decisions, oral arguments, and case decisions.

In 1992, the release of Justice Thurgood Marshall's papers and their glimpse into the Court's internal proceedings captured media attention and sparked considerable controversy (Greenhouse 1993; Weiser and Woodward 1993). Access to Marshall's papers, coupled with the availability of papers from other justices, provides judicial scholars a window on the Court's inner workings.[1] These papers contain a wealth of material, including certiorari memos, docket sheets that reveal how justices voted in conference, notices from the chief justice and occasionally the senior associate justice assigning opinions, opinion drafts, and correspondence between justices concerning the cases.

The authors gratefully acknowledge the helpful comments of Jim Spriggs, Lee Sigelman, Greg Caldeira, Harold Spaeth, and Jack Wright. The authors also thank Justice William Brennan for access to his personal papers.

[1]For a complete listing of available papers, see Epstein et al. (1994, table 5-11). Palmer and Brenner (1989, 43) provide a listing of available docket books and conference lists.

THE JOURNAL OF POLITICS, Vol. 58, No. 2, May 1996, Pp. 528–39

As one might imagine, political scientists have made increasing use of data gleaned from the justices' personal papers. Judicial scholars have used them to study certiorari (Caldeira and Wright 1994), opinion assignment (Maltzman and Wahlbeck N.d.; Cook 1992), coalition formation (Epstein et al. 1993), fluidity (Maltzman and Wahlbeck 1995; Spaeth 1994a; Dorff and Brenner 1992), and dissent behavior (Wood and Gansle 1995). At the same time, scholars have publicly and privately expressed their concerns with the reliability of the justices' papers (Ulmer 1973, 1990; Palmer and Brenner 1989; Palmer 1990), but there has been no effort to establish the reliability of the justices' records.

The absence of any attempt to empirically establish the reliability of these data is remarkable given that the justices themselves have questioned the reliability of this information. Chief Justice Warren Burger once remarked in a dispute with Justice William Douglas over the chief's opinion assignment in *Roe v. Wade* (1973) that "At the close of discussion of this case, I remarked to the Conference that there were, literally, not enough columns to mark up an accurate reflection of the voting in either the Georgia or the Texas cases" (Burger 1971). Others suggest that justices do not simply cast a vote at conference for reversal, affirmation, or another outcome, but that each justice's perceived position is distilled from their discussion of the case.[2] Thus, unlike Congress where members cast a "yea" or "nay" vote on specific propositions, justices' votes are not definitive. Even though justices use printed docket sheets with columns for affirm or reverse to summarize each justice's position, the justices' conference records are analogous to the coding of data by content analysts (see Johnson 1987).

Others have suggested that the justices' records of the conference proceedings are subject to the vagaries of the individual justices. Woodward and Armstrong (1979, 204) reported that Burger would sometimes misstate the results of the conference when recounting them to the Court's clerk. Justice Harry Blackmun would stay behind after conference to help fill in the gaps of the chief's records. Burger was undoubtedly not alone in his clerical shortcomings. Especially as justices' faculties fail them, one would expect the reliability of their records to decline. There are a number of stories of justices losing touch with their work. Most recently, Justice Marshall was accused of being disengaged, more interested in *TV Guide* than the cases on the Court's docket, and giving his clerks free reign (Eastland 1989).

While recognizing the value of information from the justices' papers, even judicial scholars who have used these papers have questioned their reliability (Palmer and Brenner 1989). As the justices record the content of conference discussions, they may err. This means that each justice's conference record may represent variations that result from "the extraneous circumstances of measurement, the hidden idiosyncracies of individual analysts" (Krippendorff 1980, 129). Without establishing the reliability of data, we have no way of knowing whether the conclusions of those scholars who have relied upon original conference data are an accurate reflec-

[2]A former clerk to a Supreme Court justice made this observation in an anonymous interview with one of the authors.

tion of Court proceedings. To establish a trustworthy basis for drawing inferences, we must determine whether the data are reliable, as the likelihood of valid conclusions is directly proportional to the reliability of the data (Krippendorff 1980). Thus, in view of the increasingly widespread use of these data, it is critical to examine the reliability of the Court's conference data.

Data and Methods

To test the reliability of the justices' records, we collected the recorded conference votes on the merits from the papers of several justices. We base our study on the 1967 and 1968 term natural court.[3] This Court is the only natural court in recent years where the docket sheets of more than one or two justices are available.[4] More specifically, we collected conference votes from the docket sheets of Earl Warren, William Douglas, William Brennan, and Thurgood Marshall. The justices' docket sheet contains a grid of justices and their votes on certiorari, jurisdiction, merits, and other motions. For the vote on the merits, which we examine here, the form has a column for each of two positions, affirm and reverse. In most cases, the recording justice simply places a check in the appropriate box to represent each justice's vote. On occasion, the justice also places a note next to a justice's vote on the merits to indicate differing grounds for the vote or to indicate a position other than affirmation or reversal.[5] Justices sometimes place additional columns in the grid to allow for votes in multiple issues or dockets.[6]

We test the reliability of the data in several different ways. First, we examine the reliability of the justices' records of votes. Taking as a given that each justice accurately assesses his own position, we compare that record with the vote recorded by other justices. If the other three justices agree with the stated position of the voting justice himself, we can conclude that their records accurately reflect the views stated at conference. We also examine the reliability of the justices' records on all votes cast at conference by all nine justices. Both the accuracy of the data and intercoder reliability are assessed by a measure of intercoder reliability. Second, we

[3]We examine records in cases that were orally argued and for which a signed opinion was published (Spaeth 1994b). We exclude votes on the cases for which the Court exercised original jurisdiction—only Justice Brennan's papers contained a completed docket sheet and the recorded positions were noted in unusual categories, for government and for state. We also exclude Warren's record of Brennan's vote in another case where Warren's record simply contains a note that is indecipherable.

[4]After the Court's 1968 term, the only justices with available docket sheets are William Douglas, William Brennan, and Thurgood Marshall (see Palmer and Brenner 1989). Thurgood Marshall's docket sheets are only available for the Court's 1967, 1968, and 1985–1990 terms.

[5]There are several other outcomes not presented on the docket sheet: dismiss as improvidently granted, dismiss, vacate, remand, modify, reargue, and hold. Justices also note reasons for no vote, like pass or out. Out is also another column on the docket sheet.

[6]One might suspect that the justice's records are later updated or corrected by his or her clerks. Of course, if this did occur, it would substantially bias the reliability of the data. However, as the docket sheets often contain written notes by the justices, it appears that the justices themselves completed the sheets. Even if the clerks do correct the sheets, though, this is not necessarily bad from the researcher's perspective as we then have more accurate data.

examine whether conference recording error is a random occurrence or whether it is more likely to occur under certain conditions. If Burger is correct, one would expect errors to occur in particularly complex cases, but one might also expect errors in recording the votes of ideological foes, justices with whom the recording justice has not served long, justices whose conference positions are tentative, and in cases where the Court lacks consensus. Moreover, one might expect more errors in the papers of a recording justice who is relatively aged. We test these expectations by estimating a logistic regression model of conference recording error.

The Reliability of Conference Data

How accurate are the justices' appraisals of their colleagues' positions? To the extent that a justice's mind is made up at conference, we have an accurate portrayal of the four recording justices' stances at conference.[7] The accuracy of the justices' records can be measured by the agreement between each justice's record of his own vote and other justices' records of that vote. We examine, for instance, the rate of agreement between Warren's record of Brennan's vote and Brennan's record of his own vote. If the two records agree, we conclude that Warren's record is an accurate depiction of Brennan's vote at conference.

To assess the rate of agreement, we use the kappa statistic. Essentially, kappa tests the null hypothesis that the observed level of agreement might occur by chance. More specifically, in the form used in this portion of our analysis, kappa is the proportional reduction of the rate of expected disagreement among coders.[8] If there is perfect agreement, kappa equals one; otherwise, the level of agreement that can be observed by chance is denoted by kappa equal to zero, while kappa less than zero indicates agreement at less than the rate expected by chance (Cohen 1960; Landis and Koch 1977).

The degree of agreement across the justices ranges from substantial, at worst, to almost perfect. The rates of agreement among Warren, Douglas, and Brennan, presented in table 1, range between 86.9% and 93.4%. Of these three justices, Justice Brennan has the highest overall rate of agreement—92.1%, 87.7%, and 93.4% agreement with Warren, Douglas, and Marshall, respectively. Uniformly, the justices' agreement rate exceed that expected by chance. Douglas has the highest rate of reduction of the expected rate of disagreement, as indicated by the kappa statistic. For two of the three justices, he has almost perfect agreement. Justice Marshall, on the other hand, has the lowest rate of agreement with his brethren, recording agreement between 80.9% and 86.0% of the votes. Nevertheless, even Marshall's records substantially agree with the voting justice's records. Thus, all four justices seem to accurately portray their fellow justices' votes.[9]

[7]Of course, the justices themselves may not have come to a firm conclusion at the conference. Indicative of this is Warren's practice to occasionally put a question mark next to his own vote.

[8]The formula for calculating kappa is $\kappa = (\pi_a - \pi_e / 1 - \pi_e)$, where π_a is the observed rate of agreement and π_e is the expected rate of agreement (Landis and Koch 1977, 163).

[9]A problem, however, arises as the voting categories that justices use on their docket sheets are not mutually exclusive, which reduces reliability. To resolve this problem, if one recording justice

TABLE 1

ACCURACY OF RECORDED CONFERENCE VOTES

Recording Justice	Voting Justice	Rate of Agreement	Kappa
Warren[a]	Douglas	86.89%	.7622***
	Brennan	92.08%	.8142***
	Marshall	90.05%	.8518***
Douglas[a]	Warren	88.35%	.7422***
	Brennan	91.13%	.8007***
	Marshall	89.05%	.8369***
Brennan[b]	Warren	92.12%	.8034***
	Douglas	87.68%	.7723***
	Marshall	93.43%	.9003***
Marshall[c]	Warren	82.89%	.6350***
	Douglas	80.92%	.6353***
	Brennan	86.00%	.7282***

Note: [a]$n = 206$; [b]$n = 203$; [c]$n = 152$; ***$p < .0001$.

Given the significant pairwise agreement on votes cast at conference, do the docket sheets of Warren, Douglas, Brennan, and Marshall record the same vote for all nine justices? We can assess this by examining the votes of all nine justices recorded by our four justices. Again, we use kappa as a measure of intercoder reliability. In this form, kappa compares the number of coders who associate a justice with a particular outcome and the total number of coders. The kappa statistic is then calculated for each outcome, like affirm, reverse, and so on, and for the combined outcomes.[10] This form of kappa has the advantage of accounting for multiple coders and multiple outcomes.

Two outcomes are explicitly presented on the printed docket sheet, affirm and reverse. Not surprisingly, there is a high degree of intercoder agreement for these

described a justice's vote as a pass, out, or dismissed as improvidently granted, we recoded votes originally recorded *no vote* to *pass* or *out* and recoded *dismiss* to *dismiss as improvidently granted*. Predictably, the rates of agreement increased as a result of this recoding. The most marked increase occurred with Warren's, Brennan's, and Marshall's coding of Douglas' vote. The agreement rate increases, respectively, to 94.17%, 96.55%, and 85.53%. This occurs because Douglas frequently passed or did not participate. Kappa also improves, as one might expect, from this recoding. Warren and Douglas now each have almost perfect agreement with at least two of the other justices and Brennan has almost perfect agreement with all three voting justices.

[10]The formula for calculating kappa with multiple outcomes, $k \geq 2$, and multiple coders is

$$\kappa = \frac{\sum_{j=1}^{k} \bar{p}_j \bar{q}_j \kappa_j}{\sum_{j=1}^{k} \bar{p}_j \bar{q}_j}$$

where $\bar{p}_j$ is the overall proportion of ratings in category j, $\bar{q}_j = 1 - \bar{p}_j$, and κ_j is the value of kappa for category j (Fleiss 1981, 229).

TABLE 2

INTERCODER RELIABILITY OF CONFERENCE VOTE ON MERIT

Outcome	Kappa
Affirm	.8747
Reverse	.8462
Affirm in part/Reverse in part	.2183
No vote	.1859
Pass	.3988
Out	.9036
Dismiss	.3110
DIG	.3844
Remand	.4785
Vacate	.5269
Modify	.1181
Reargue	.0719
Hold	.6665
Combined	.7706
n = 1,852	

two outcomes.[11] Indeed, as reflected in table 2, there is almost perfect agreement when the outcome takes one of these two values as the kappa statistic is, respectively, .8747 and .8462. The only other outcome that has this degree of reliability is when the justices did not participate in the case (kappa = .9036). Other than when the voting justice prefers to hold the case over, there is only moderate agreement, at best, on the placement of votes for other outcomes. There is only slight or fair agreement among the justices on votes for dismissal, dismiss as improvidently granted, modify, reargument, and affirm in part/reverse in part. There is, similarly, a lack of agreement on passes and no recorded votes. Intuitively, one would expect greater disagreement for these outcomes since, in order to record a vote for an alternative outcome, the justice has to either make a note on the sheet or alter the heading on the column from, say, reverse to remand (see Holsti 1969; Krippendorff 1980). Nevertheless, as seen in the kappa for all outcomes combined, there is substantial agreement among the recording justices on the outcome supported by each of the nine justices.[12]

While the justices' records are reliable—that is, they significantly depart from chance agreement—one might suspect that the unreliability exhibited in the data,

[11]For this analysis, we exclude the records of the justice who voted if he is one of the four recording justices. This is done to prevent biasing the reliability results since presumably if we used the records of the voting justice, those will be accurate.

[12]Unfortunately, when the number of coders varies across cases, as happens when justices do not have a vote recorded because of absence or confusion, we technically cannot test the null hypothesis of chance agreement. Some have argued that showing that the estimate differs from zero does not have much meaning, given that we are mainly concerned with how much above zero it is (Cohen 1968, 217).

particularly in Thurgood Marshall's papers and in votes other than to affirm or reverse, has important substantive implications. This can be examined by exploring fluidity in votes on the merits—the extent to which justices change their votes between conference and the release of the final decision. If these data are substantially unreliable, one would expect this to be reflected in significant differences in the rates of fluidity reported for each of the nine voting justices across the four recording justices. To assess this substantive impact, we first calculate a baseline level of fluidity for each sitting justice by determining the conference vote that the majority of recording justices attribute to a justice. Then, by comparing this vote with the justice's final vote on the merits, as reported in Spaeth (1994b), we ascertain the presence of fluidity.[13] Using a difference in means test, we are able to determine whether rates of fluidity systematically vary between each recording justice and the majority consensus.

There is not a significant substantive difference in fluidity reported by different recording justices. When we compare the fluidity rates assigned to each of the nine voting justices by the four recording justices and the fluidity rates assigned to the nine justices using the conference votes supported by the majority of recording justices, no recording justice, not even Justice Marshall, reports fluidity that is statistically distinguishable from the majority consensus. Moreover, when we examine fluidity in the generally less reliable conference votes, that is, votes other than to affirm or reverse, no recording justice's fluidity rate for the Court is significantly different from that calculated by the majority. From these tests, we conclude that the data are not only statistically reliable, but are also substantively reliable.

Model of Justice Recording Error

Are the justices' errors random occurrences or can they be explained as a product of certain conditions? As Chief Justice Burger asserted in his exchange with Justice Douglas, justices might be more prone to error in complex cases. When justices discuss and vote on multiple issues or on several cases consolidated on appeal, one justice may record the vote on one issue or case, while a second justice records the vote on another. Second, one might suspect that a justice is most likely to know the mind of a justice who is an ideological ally, rather than an ideological foe; this might result from more frequent informal communications on cases. Third, justices may learn the inclinations and views of justices as they serve together on the

Nevertheless, to test the null hypothesis of chance agreement, we calculate the kappa statistic for observations on which all four justices had recorded the conference vote. The combined kappa was .7863. This was statistically different from chance agreement at less than .0001. The kappa statistic for each outcome is almost identical to that reported in table 2. For each outcome, except hold, the rate of agreement departs from chance occurrence at a statistically significant level, $p < .0001$.

[13]We limit our analysis to conference votes to affirm, reverse, remand, vacate, dismiss, and partial affirm and partial reverse. We do this because these votes have analogues in the final vote on the merits reported by Spaeth (1994b).

Court. Thus, we speculate that justices will make more mistakes in recording votes of those justices with whom they have not served very long. Fourth, one might expect justices who are relatively old to make more recording errors than other justices. Fifth, if the voting justice's position at conference is tentative, we can expect error. Finally, we expect justices to be less apt to err in cases where there is a consensus on the Court. If most justices agree on the case outcomes, it is not as likely that a justice's position will be misinterpreted by his or her colleagues.

To test these hypotheses, we use a dichotomous dependent variable that identifies whether a recording justice disagreed with the majority of recording justices in assessing a justice's vote.[14] We operationalized the independent variables in the following ways. Multiple Issues identifies cases that involve multiple issues according to the United States Supreme Court Judicial Database (Spaeth 1994b).[15] Multiple Dockets similarly indicates whether a decision contains more than one docketed case (Spaeth 1994b). Ideology is the percentage of cases in that issue on which the voting justice and the recording justice voted together (Epstein et al. 1994, table 6-3).[16] Joint Service is the number of full terms that the voting justice and the recording justice have served together on the Court (Epstein et al. 1994, table 4-12).[17] Age equals the recording justice's age in October, the month when the Supreme Court's term begins (Epstein et al. 1994, table 4-2). Tentative Vote is a dichotomous variable that equals one if the recording justice put a question mark next to the recorded vote. Finally, Consensus equals one if the original conference vote was unanimous or had only one vote in the minority.

The logistic regression, presented in table 3, suggests that the conference recording error is more than just random noise. The model chi-square is statistically significant, so we must reject the null hypothesis that the estimates of all our variables are equal to zero. What is more, many of our independent variables are statistically significant. Chief Justice Burger's hypothesis that error is likely to occur in complex cases is supported, at least with respect to cases with multiple dockets. If a decision contains multiple cases, the likelihood of recording error increases from 4.0 percent to 6.0 percent.[18] It is a testament to the overall accuracy of these data

[14] An observation, for our purposes, is one recording justice's record of a single justice's vote. Between the four recording justices and nine voting justices, there was a total of 6,137 recorded votes in 209 cases that were orally argued and formally decided. We excluded the records of 176 votes for which the recording justices had no clear consensus. We also excluded the justice's record if he was the voting justice.

[15] More specifically, we sorted all decisions from the 1967 and 1968 terms of the Court by U.S. Reports citation to identify all unique decisions. If such a case had more than one issue reported, using the narrow issue areas, Multiple Issues equals one.

[16] If a case involved more than one issue, we use the issue on which the justices voted together most frequently.

[17] The number of full terms that two justices have served together is calculated from the date of each justice's confirmation.

[18] To calculate the probability of recording error, we set the value of each independent variable at its mean and allow the value of one variable to vary.

TABLE 3

LOGISTIC REGRESSION OF CONFERENCE RECORDING ERROR

Variable	Estimate (Standard Error)
Constant	−1.551*
	(.773)
Multiple issues	.245
	(.192)
Multiple dockets	.427*
	(.190)
Ideology	−.003
	(.005)
Joint service	−.054***
	(.013)
Age	−.016
	(.010)
Tentative vote	1.159**
	(.485)
Consensus	.562
	(.218)
Number of observations	5,821
Log likelihood	−1,051.0421
Chi-square	47.01 with 7 d.f.***

*$p < .05$ one-tail test; **$p < .01$ one-tail test; ***$p < .001$ one-tail test.

that a substantively small increase is significant. Cases with multiple issues, on the other hand, do not predispose justices to error.[19]

Much of judicial behavior gravitates around the justices' policy views, but although justices vote together, they are not more apt to accurately assess their allies' positions than their ideological foes. The coefficient for ideology is negative, indicating that justices are more likely to err in recording the votes of justices with whom they do not typically agree, but ideology is statistically insignificant. One might suspect that ideology might be captured indirectly through other independent variables, but there is no evidence of multicollinearity.[20]

The recording justice's experience and length of service with colleagues enhances the accuracy of their records. A justice who was recording the vote of another justice with whom he served together for 29 terms has a likelihood of recording error of only 1.3%. In contrast, a justice who was recording the vote of a colleague with whom they had just starting serving has a 6.0% likelihood of erring.

[19]Further analysis reveals that Multiple Issue and Multiple Docket are not collinear. The correlation coefficient for these two variables is −.07.

[20]To test for multicollinearity, we regressed ideology on the other independent variables. The resulting R-square was about .11. This is sufficiently low to reject the possibility of multicollinearity.

On the other hand, age is not significantly related to the odds of making recording mistakes. This finding may reflect a nonlinear relationship between age and recording error propensity—age may only become an important factor when the justice is so advanced that he is effectively disabled.

The nature of the justice's position at conference affects the accuracy of the justices' records. If the voting justice expressed a somewhat tentative stance, as reflected in question marks next to the vote, the records are more likely to be in error. The recorded vote of a justice who is tentative is 8.0% more likely to be in error than the vote of a more certain justice. The nature of the majority coalition also affects the accuracy of docket sheets. Justices, somewhat surprisingly, are more likely to make mistakes in cases that enjoy consensus on the Court. The coefficient for Consensus is positive, indicating a positive relationship between consensus and recording error.

Conclusion

To assess the reliability of conference votes, we examined the docket sheets in the personal papers of Justices Warren, Douglas, Brennan, and Marshall. By comparing each justice's record of his own vote with the records of other justices, we were able to assess the accuracy of the recording justices' docket sheets. The recording justices agreed with the voting justice between 81% and 93% of the time. The most accurate justices were Brennan and Douglas, while Marshall lagged behind the other justices. More importantly, all four justices' records significantly departed from mere chance agreement with the voting justice. We also found substantial agreement among the four recording justices on the positions expressed by all nine justices. Specifically, we found almost perfect agreement when the voting justice cast a vote for either affirm or reverse; agreement among the recording justices was moderate if the other justice cast a vote for another outcome.

What are the implications of this study? First, although the justices' conference records are not perfect, these data are reliable portrayals of the justices' expressed positions at conference. The most accurate records are those of Brennan and Douglas. Although Marshall's records are reliable, they are less reliable than docket sheets contained in the other justices' papers. If a researcher has a choice, which presently is not true for terms after 1985, the researcher should use another justice's papers. Second, justices' records are particularly reliable when the voting justice supports affirmation or reversal. On the other hand, if the recorded vote supports another outcome, the reliability of the data decreases, especially for those outcomes that are not mutually exclusive. Thus, when the recording justice reports that another justice supported a dismissal, dismissal as improvidently granted, modification of the lower court's order, or a reargument, the researcher should use those data with caution. Finally, since the logit model of conference recording error finds systematic error in the justices' records, scholars using the justices' papers should, when possible, take cautionary steps. For instance, scholars may

examine multiple justices' papers in cases involving multiple dockets, tentative votes, and the records or votes of newcomers to the Court.

Manuscript submitted 15 May 1995.
Final manuscript received 18 August 1995.

References

Burger, Warren. 1971. "Letter to William Douglas, December 20, 1971." Available in Justice William O. Douglas' papers. Washington, DC: Library of Congress.

Caldeira, Gregory, and John R. Wright. 1994. "Nine Little Law Firms? Justices, Organized Interests, and Agenda-Setting in the Supreme Court." Presented at the annual meeting of the Midwest Political Science Association, Chicago.

Cohen, Jacob. 1960. "A Coefficient of Agreement for Nominal Scales." *Educational and Psychological Measurement* 20:37–46.

Cohen, Jacob. 1968. "Weighted Kappa: Nominal Scale Agreement with Provision for Scaled Disagreement or Partial Credit." *Psychological Bulletin* 70:213–20.

Cook, Beverly B. 1992. "Testing a Model of Opinion Assignment: The Burger Court." Presented at the annual meeting of the American Political Science Association, Chicago.

Dorff, Robert H., and Saul Brenner. 1992. "Conformity Voting on the United States Supreme Court." *Journal of Politics* 54:762–75.

Eastland, Terry. 1989. "While Justice Sleeps: Some of our Justices Are Old and Sick, Others Have Simply Relinquished Power to their Clerks." *National Review* 41:24–26.

Epstein, Lee, Carol Mershon, Jeffrey A. Segal, and Harold J. Spaeth. 1993. "Collaborative Research on the Formation of Opinion Coalitions on the U.S. Supreme Court." National Science Foundation Grant Proposal.

Epstein, Lee, Jeffrey A. Segal, Harold J. Spaeth, and Thomas G. Walker. 1994. *The Supreme Court Compendium: Data, Decisions, and Developments.* Washington, DC: Congressional Quarterly Press.

Fleiss, Joseph L. 1981. *Statistical Methods for Rates and Proportions.* 2d ed. New York: Wiley.

Greenhouse, Linda. 1993. "Justices Guard Mystique." *The New York Times*, 26 May, sec A.

Holsti, Ole. 1969. *Content Analysis for the Social Sciences and Humanities.* Reading, MA: Addison-Wesley.

Johnson, Charles A. 1987. "Content-Analytic Techniques and Judicial Research." *American Politics Quarterly* 15:169–97.

Krippendorff, Klaus. 1980. *Content Analysis: An Introduction to Its Methodology.* Beverly Hills: Sage.

Landis, J. Richard, and Gary G. Koch. 1977. "The Measurement of Observer Agreement for Categorical Data." *Biometrics* 33:159–74.

Maltzman, Forrest, and Paul J. Wahlbeck, N.d. "May It Please the Chief? Opinion Assignment in the Rehnquist Court." *American Journal of Political Science.* Forthcoming.

Maltzman, Forrest, and Paul J. Wahlbeck. 1995. "To Affirm or Reverse? Strategic Policy Considerations and Voting Fluidity on the Burger Court." Presented at the annual meeting of the Midwest Political Science Association, Chicago.

Palmer, Jan. 1990. *The Vinson Court Era: The Supreme Court's Conference Votes.* New York: AMS Press.

Palmer, Jan, and Saul Brenner. 1989. "Working with Supreme Court Docket Books." *Law Library Journal* 81:41–46.

Rehnquist, William H. 1987. *The Supreme Court: How It Was, How It Is.* New York: Quill.

Roe v. Wade. 1973. 410 U.S. 113.

Spaeth, Harold J. 1994a. "The Breakup of the Original Vote Coalition on the Vinson Court: Is the Conference Vote on the Merits Really the Final Vote?" Presented at the annual meeting of the Southern Political Science Association, Atlanta.

Spaeth, Harold J. 1994b. *United States Supreme Court Judicial Database, 1953–1992 Terms.* Ann Arbor: Inter-University Consortium for Political and Social Research.

Ulmer, S. Sidney. 1973. "Bricolage and Assorted Thoughts on Working in the Papers of Supreme Court Justices." *Journal of Politics* 35:286–310.

Ulmer, S. Sidney. 1990. "Further Reflections on Working in the Papers of Supreme Court Justices." *Judicature* 73:193–96.

Weiser, Benjamin, and Bob Woodward. 1993. "Roe's Eleventh-Hour Reprieve." *Washington Post*, 23 May, sec A.

Wood, Sandra L., and Gary Gansle. 1995. "Mustering the Minority: The Dissent Assignments of William J. Brennan Jr." Presented at the annual meeting of the Midwest Political Science Association, Chicago.

Woodward, Bob, and Scott Armstrong. 1979. *The Brethren: Inside the Supreme Court.* New York: Avon Books.

Forrest Maltzman and Paul J. Wahlbeck are assistant professors of political science, George Washington University.