

---

## Understanding Judicial Hierarchy: Reversals and the Behavior of Intermediate Appellate Judges

---

Kevin M. Scott

One of the central controversies in the judicial behavior literature is the extent to which judges' ability to act according to their ideological preferences is affected by their location in the judicial hierarchy. Judges on intermediate appellate courts have different decisionmaking environments than high court judges. As a result, the goals of lower appellate court judges may differ from those of their superiors: the quest for legal accuracy may compete with the desire to pursue policy preferences. Analysis of the reversal rate of the U.S. circuit courts of appeals offers insight into the extent to which these judges balance the pressures of their own policy preferences with the desire to achieve the legally accurate result in cases they decide.

**I**nteraction between first-level and second-level appellate courts is a subject that has garnered increasing attention over the past several years, but an opportunity to bridge disciplinary divides by integrating recent scholarly work by political scientists and legal academics has largely been missed in the process. Disagreements between the two levels of courts produce reversals of the lower court by the higher court, but the process that drives such reversal is not well-understood given its importance in maintenance of judicial organization. Particularly in systems where review by the second-level appellate court is discretionary, higher courts reverse lower courts in pursuit of making broader policy statements, so the process takes on added importance. Work that assesses decisions by high courts to grant leave to appeal (Atkins 1990; Caldeira et al. 1999; Cameron et al. 2000; Flemming 2004; Flemming & Krutz 2002; Perry 1991; Provine 1980) offers important insight into how second-level courts select their cases. But the factors that determine courts' decisions to review each case may not offer a complete pic-

---

This article was originally presented at the 2004 Annual Meeting of the Midwest Political Science Association. I would like to thank Christopher Banks, Larry Baum, Chris Bonneau, Rhonda Evans Case, Lori Hausegger, Micheal Giles, Steve Wasby, and Margaret Williams for their comments and advice. I would also like to thank Michael Bailey for his advice on data analysis. Please address correspondence to Kevin Scott, Department of Political Science, Texas Tech University, Box 41015, Lubbock, TX 79409-1015; e-mail: kevin.scott@ttu.edu.

ture of the relationship between first- and second-level appellate courts. Assessing the determinants of a lower court's reversal rate offers an opportunity to develop a fuller understanding of the process of review and reversal. In doing so, it is possible to more completely develop a model of what influences the behavior of intermediate appellate judges.

Scholars who address the phenomenon of judicialization of politics (Guarnieri & Pederzoli 2002; Ginsburg 2003; Hirschl 2004; Stone Sweet 2000; Shapiro & Stone Sweet 2002; Tate & Vallinder 1995) note that the importance of courts and judges has arisen alongside the belief that judges are political actors, motivated by their own views about what constitutes good public policy. But viewing judges as (mostly) political actors may not apply as well to judges on lower courts. To a substantially greater degree than high court judges, lower appellate judges may be motivated by a desire to make good law as well as a desire to make policy. This behavior may be expected for at least four reasons. First, lower court judges may simply seek to maximize efficiency: consistent law, even though it may contradict judges' policy preferences, makes judging easier and increases the consistency of results for litigants. Second, higher court judges hear more hard cases, those not easily settled by existing law. Third, the higher one moves in the judiciary, the less the vertical doctrine of precedent applies (Baum 1997; Greene et al. 1998). Fourth, judges on the lower courts may simply desire promotion to the high court. Although strategies for pursuing this may vary, one way to become noticed is to develop a reputation for judicial quality, including being reversed rarely, if ever, by the high court (Abe 1995; Salzberger & Fenn 1999; but see Klein & Hume 2003). All of this means that judges on intermediate appellate courts are more constrained by decisions that emanate from the courts above them, limiting the cases in which they can pursue their policy preferences and the range of options they have when they are free to do so.

The belief that judges make good-faith efforts to correctly interpret and apply the law independent of their policy preferences suggests that ideological disagreement between higher and lower courts would not affect variation in the reversal rate within or across lower courts. Rather, the ability of the judges of the lower court to correctly apply the law would determine how frequently a given judge or court is reversed. If components of the court, particularly its administrative structure, hamper this activity, then that court would experience a greater reversal rate than other courts at the same level of the hierarchy.

Neither a model that views intermediate appellate judges as motivated by good law nor a model that focuses on their policy preferences may be completely accurate as applied to judges on

intermediate-level appellate courts. The two approaches, though, should be integrated to test the extent to which they can help explain the relationship between first- and second-level appellate courts. There is room in our studies of judicial behavior for both models, and hopefully incorporating the two will provide a more complete version of the field of forces that influence the behavior of appellate judges.

## Theoretical Issues

Literature positing that appellate judges are motivated by the desire to make good law takes little notice of arguments that judges are motivated by their policy preferences, but the reverse is also true. Cross and Tiller, for example, contend that “much of the [legal] scholarship simply assumes the sincere application of legal doctrine without considering the possibility that it may at times be nothing more than a convenient rationalization for political decision-making” (1998:2156). The same criticism may be made of the work of political scientists—that they give inadequate attention to the role of law in the decisions of high court judges, focusing almost exclusively on the policy preferences of the judges (Segal & Spaeth 2002). But, as Baum (1997) argues, finding that high court judges rely on policy preferences to reach decisions does not foreclose the possibility that judges at lower levels of the judiciary use the law as a guide.

Finding empirical evidence of the pursuit of legal goals has long troubled those who study the behavior of appellate judges. Much of this stems from two sources. First, as Cross and Tiller indicate, the belief that judges give primacy to accurate application of doctrine is more an assumption than a proven fact (1998). Second, scholars have struggled to develop testable hypotheses that would demonstrate that judges do endeavor to follow the law in their behavior. At the lower court level, though, a series of studies has indicated that judges adjust their behavior to accommodate changes in higher courts (Songer et al. 1994; Cameron et al. 2000). This could occur for two reasons. First, lower court judges could be strategic actors who recognize that following preference shifts that occur further up the judicial hierarchy minimizes the chance of reversal and maximizes their impact on policy outcomes. Second, as membership shifts on the higher court, so does its doctrine, and lower court judges may be responding to the shifts in doctrine. The conflicting possibilities highlight the problem of developing hypotheses to test if judges at any level actually attempt to make good law.

One way to address this problem may be to look at how frequently lower court judges are reversed by judges at other levels. In this context, one can argue that judges seek to make good law, but some feature of their environment may affect their ability to do so. For example, courts that are larger than other courts at the same level in the hierarchy may be reversed more frequently, not just because they generate more decisions that are candidates for reversal (that is, the number may be higher, but the rate may be the same), but because the quality of their work is affected by their environment.<sup>1</sup> But if one assumes that appellate judges are motivated by their policy preferences more than by their desire to make good law, then one would expect that the primary predictor of the frequency with which a lower court is reversed is a function of its ideological distance from the high court.

### Existing Research

Scholars know surprisingly little about how superior courts deal with intermediate appellate courts. Several system-level phenomena are easily observed: for supreme courts with discretionary jurisdiction, for example, one can assess the frequency with which lower court decisions are reviewed by the high court. Variation in the review rate provides some insight into the relative power (or political importance) of supreme courts and the intermediate appellate courts below them. High courts that hear fewer cases may be less influential than those that hear more cases, at least relative to the other courts in the hierarchy. Another easily observed indicator is the frequency with which high courts intervene in the decisions of intermediate appellate courts. Generally speaking, the lower the intervention rate, the more respected the intermediate appellate courts. But comparison of this data across systems may mask important internal variation that may provide important insight into the nature of the relationship between the two levels of the judiciary. If there is substantial variation in the rate at which a higher court reviews or reverses its lower courts (or, in systems with unified intermediate courts, the judges of the lower court), then understanding the sources of that variation may provide insight into the nature of decisionmaking and judicial power.

Work on the role of high courts, much of which documents a trend toward an increasingly important role for the courts in resolving political disputes (evolving from private to public law), unfortunately pays little attention to the role of intermediate courts.

---

<sup>1</sup> In one respect, the hypotheses developed here could generally be tested at the level of the judge, but institutional factors (court size, workload) and some ideological factors (heterogeneity) are court-level dynamics, affecting all judges on a court equally.

In civil law systems, constitutional courts often operate independently of the judicial hierarchy, serving more as a complement to the remainder of the national government, while the judiciary is topped by a supreme appellate court that is not a focal part of the national political debate (for example, France's Constitutional Council and Supreme Court of Appeal), so the question of the politics of judicial hierarchy is separate from questions of constitutional interpretation. In most common-law systems, the judiciary has operated at the outer perimeter of the political process, so questions about the political impact of decisions are relatively new (Epp 1998; McCormick 2000; Pierce 2002), and study of intermediate appellate courts lags behind those works.

Atkins (1990), in his study of the reversal rates of the Court of Appeal in Britain and the U.S. courts of appeals, argues that one might expect the U.S. Supreme Court to reverse more lower court decisions (reverse a greater percentage of the cases they review) than do the Law Lords. Some of the reasons for this expectation are systemic: the centralization of the English courts, their status relative to Parliament, and the relative homogeneity of the judiciary all serve to limit the amount of conflict between the two levels relative to what would be expected in the American federal judiciary. But cases that filter up from trial courts to the Court of Appeal are more likely to present novel issues (issues that present an unsettled question of law) than those that move up from the U.S. district courts, so one might expect closer supervision of the Court of Appeal than of decisions made by the U.S. courts of appeals (Atkins 1990).

Atkins's central findings, that both the Court of Appeal and the U.S. courts of appeals are the final voice on a vast majority of cases, and that the U.S. Supreme Court reverses a much higher percentage of the cases it hears than do the Law Lords, offer several further opportunities for study. Perhaps most important, there is a need for a better understanding about the internal dynamics of court hierarchies. In the English system, this promising avenue of research has not yet been fully developed. An important exception is Salzberger and Fenn (1999), who look at the relationship between support for the government by justices on the Court of Appeal and promotion to the Judicial Committee of the House of Lords. They find that Conservative governments do not promote pro-government justices more quickly. They do find, however, that pro-government justices who reverse fewer lower court decisions are promoted more quickly (1999:846).<sup>2</sup> Perhaps more relevant to

---

<sup>2</sup> There were insufficient data to reach any conclusions on Labour governments. In addition, Salzberger and Fenn do not test whether judges support the government only when one party is in power.

the current question, there appears to be a fair amount of variation in support for the government and in the reversal rate among Court of Appeal justices, suggesting that some leverage may be gained in studying the determinants of that variation.

Somewhat more work has been done to determine the sources of variation in reversal rates of systems where intermediate appellate courts are more fractured than the English Court of Appeal. The Canadian provincial courts of appeal (PCAs), for example, share features of both the American and British systems. Like the U.S. courts of appeals, they are arranged geographically, and vary considerably in size and political composition (Greene et al. 1998; McCormick 1992). Like the Court of Appeal, judges of the PCAs are appointed by Ottawa. Perhaps because of the shared source of appointment for Canadian Supreme Court justices and PCA judges, scholars have struggled to ascertain clear patterns in the levels of attention given to the different courts of appeal. For example, McCormick (1992) finds little relationship between court size (and province population) and number of reviews of PCA decisions. This finding is confirmed by Flemming and Krutz (2002), who find no evidence of geographical variation in the success of leave applications from the different courts of appeal, including the Federal Court of Canada (see also Flemming 2004). Importantly, in both the Canadian and British contexts, very little work has been done to assess the relationship between ideological disagreement and reversal frequency. McCormick (1992) makes an initial attempt along these lines, finding that, across the different provincial courts of appeal in Canada, Liberal appointees are reversed more frequently than Conservative appointees, but one must also consider the composition of the Canadian Supreme Court before assessing this finding.

One might expect that, given the importance of the U.S. Supreme Court to American politics, literature on the relationship between the Supreme Court and the courts of appeals would be quite well developed. But surprisingly little work has looked at variation in the success rates of the different U.S. courts of appeals before the Supreme Court. The study that most clearly addresses this question (though it looks at summary reversals, not full reversals<sup>3</sup>) finds that circuit size plays a role (Posner 2000) but does not account for all the variation in the rate at which circuits are reversed summarily by the Supreme Court. Posner's work, however, suffers an important shortcoming. As argued above, if one assumes that appellate judges might behave according to their

---

<sup>3</sup> This occurs when the Supreme Court does not hear arguments or write a full opinion, usually suggesting that the lower court was so clearly mistaken in its decision that the Supreme Court need not fully consider the case.

policy preferences, then any assessment of the predictors of reversals should measure the ideological distance between the high court and the intermediate appellate court(s).

Several commentators and a few scholars have focused their attention on an interesting phenomenon in the American federal judiciary. The Ninth Circuit Court of Appeals, which covers much of the western United States, is reversed far more frequently than any other circuit. Its critics have pointed to its size (Posner 2000), but others have noted its substantial ideological distance from the Supreme Court, particularly over the last 25 years (Hellman 2000; Herald 1998; Wasby 1998). Another potential source of the Ninth Circuit's reversal rate is that it, unlike any other circuit, uses a limited en banc procedure to review panel decisions, while all other circuits use full en banc proceedings to accomplish the same purpose.<sup>4</sup> In testimony submitted to the Commission on Structural Alternatives for the Federal Courts of Appeals, Justice Antonin Scalia identifies the limited en banc proceeding as the source of many of the Ninth Circuit's problems. He notes that a "disproportionate segment of the [Supreme] Court's docket . . . is consistently devoted to reviewing Ninth Circuit judgments, and to reversing them by lop-sided margins, [which] suggests that the [limited en banc] error reduction function is not being performed effectively" (Scalia 1998). But Wasby finds that this does not mean that the Supreme Court treats cases reviewed by the limited en banc with any more deference than other Ninth Circuit cases, and, more generally, that "there is little evidence that the Supreme Court, in either its outcomes or opinions, cuts the United States Courts of Appeals much slack for having decided cases en banc" (2001:73).

While the relative role of ideology and law (and the evidence collected to support them as determinants of appellate decision-making) may vary across systems, there appears to be a consistent desire among intermediate appellate court judges to be promoted. Salzberger and Fenn find that judges on the English Court of Appeal who are of higher perceived quality are promoted more quickly, suggesting that judges interested in promotion may take care to develop a record of quality (1999). McCormick (1992) notes that most justices on the Supreme Court of Canada are former

---

<sup>4</sup> On all circuits, cases are normally decided by three-judge panels. On every circuit but the Ninth, a request from a judge or litigant to review a case en banc is reviewed by all of the judges and, if a majority of those voting (all active judges, as well as senior and visiting judges from the original panel) agree, all the judges in the circuit rehear the case. The Ninth Circuit hears cases en banc with only 11 of the active judges participating (10 judges chosen at random and the chief judge). There are currently 28 authorized judgeships for the Ninth Circuit. While fewer than half hear the case en banc, the decision to rehear a case en banc in the Ninth Circuit is handled the same way as in other circuits.

court of appeal judges, and Epstein, Knight, and Martin (2003) point to an increasing norm of prior judicial experience for U.S. Supreme Court justices. The chance for promotion, while varying across countries, is likely sufficient to induce behavior by lower court judges that they view as enhancing their chances for promotion. Such a desire likely does not vary among judges within a country. The effect of this desire should be to encourage judges to strive to minimize reversal of their decisions by a higher court.<sup>5</sup>

A model of decisionmaking for appellate judges, then, should leave ample room for both ideological and legal goals. Cross, for example, finds that both legal and ideological criteria influence the behavior of judges on the U.S. courts of appeals (2003; see also Klein 2002). Judges may pursue both their policy preferences and good law; they may vary in which objective they pursue across cases or policy issues. In considering the collective behavior of the judges, some judges may be more committed to pursuing their policy goals while others may be more interested in the pursuit of legal accuracy. The two goals may frequently produce the same result. But they may also produce divergent results, and the patterns of higher court reversal of lower courts may help an understanding of the degree to which the two competing models of behavior of intermediate appellate judges accurately reflect reality. Understanding the variation across circuits in reversal rates will offer some important insights into the factors that prove relevant to the behavior of intermediate appellate judges.

## Hypotheses

I test the mechanics of judicial hierarchy in the context of the American federal judiciary. The particular appeal of this context is that the geographical organization of the circuits allows assessment of the U.S. Supreme Court's relationship with the different circuits. The variation among the circuits is also considerable. Because judges on the lower federal courts are appointed under senatorial courtesy (the president consults home-state senators of his own party when filling a vacancy in that state), the ideological variation in lower court judges can be quite great.<sup>6</sup> In addition, the circuits

---

<sup>5</sup> A desire for promotion may also produce more ideological behavior. Given the political control of judicial promotion in common-law countries, appeals judges may seek to distinguish themselves by their ideological record rather than a reputation for quality. Different judges likely pursue different strategies to attract the attention of those responsible for selecting judges for promotion.

<sup>6</sup> Though each circuit spans several states, seats on them are allocated to the states in rough proportion to population. President George W. Bush was criticized by Maryland's (Democratic) senators when he nominated a lawyer from Virginia to replace a judge from Maryland who retired (Mikulski 2003).

vary considerably in terms of size and workload, as well as how they fare before the Court. The intercircuit variation in success rates before the Court offers interesting insight into the nature of judicial hierarchy more generally.

The proposed division between ideological and legal factors is well-developed in the literature on the American federal judiciary. Students of the U.S. Supreme Court have long argued that the institutional features of the Court permit the justices to pursue their policy preferences, and judicial review of legislation gives the Court an important role in the American political system. The belief that judges are motivated by their policy preferences is nearly as well embedded in the study of U.S. appellate court judges (Goldman 1966, 1975; more recently, Cameron et al. 2000; Hettlinger et al. 2004) as it is in the study of Supreme Court justices. But the literature on court of appeals judges does not appear to make the same claim some have made with respect to Supreme Court justices—that they “engage, in almost all cases, in rationally sincere behavior” (Segal & Spaeth 2002:350). Indeed, it appears clear that other factors influence the behavior of court of appeals judges—their colleagues (Cross & Tiller 1998; Van Winkle 1996) and the Supreme Court (Cameron et al. 2000)—which attenuate the relationship between their ideological preferences and their behavior.

Factors that enhance or interfere with judges’ attempts to make good law as well as ideological characteristics of the judges and the superior court should determine the rate at which a higher court reverses a lower court. In the American context, the size of the circuit on which the judge sits and the caseload of the judges appear to be the two issues that would most affect judges’ efforts to make good law. At the most basic level, more judges (and more panels) mean more decisions that could be reversed by the higher court. Such an argument would be relevant if one were measuring the raw number of reversals rather than the reversal rate (the percentage of cases decided that are reversed). When looking at the reversal rate of the circuit, the argument is slightly different. Simply put, as circuit size increases, mistakes should be more frequent. This may happen for several reasons. First, the more judges on a court, the weaker the operation of “informal norms of judicial propriety and restraint” (Posner 2000:712). In effect, group dynamics that force adherence to internal norms work better on smaller circuits. Second, and related, it is easier for circuit preference outliers to diverge from circuit preferences and get away with doing so (Posner 2000). This may happen because there are more attempts to defect from the wishes of the circuit median, and because the circuit median would be less likely to catch such attempts. Third, the larger a circuit becomes, the less familiar judges

become with their coworkers, as they are less likely to be assigned to panels together. If judges work together as “resource-constrained teams” (Kornhauser 1995), their ability to develop successful working relationships (and to produce correct decisions) will be affected by this infrequency of shared panel assignments. Fourth, large circuit size can increase the number of decisions that serve as binding precedent on panels. Though lawyers certainly bear the responsibility for locating precedent favorable to their clients in an adversarial system of justice, the production of decisions in large circuits can be less orderly than in small circuits. This suggests that as the size of a circuit increases, the rate at which incorrect decisions are made will increase correspondingly. As a result, there should be a positive relationship between a circuit’s size and its reversal rate.

*Hypothesis 1: The larger a circuit is, the higher its reversal rate will be.*

This argument is not without complications. As it rests on the assumption that judges attempt to make good law, it is difficult to extricate the rationale of this argument from the notion that the Supreme Court reverses decisions it finds ideologically incompatible. A modification of this hypothesis may solve this problem. If one assumes that *unanimous* Supreme Court decisions are those that represent non-ideologically driven error correction of lower court decisions, then it may be the case that the larger a circuit grows, the more problems it will have monitoring panel decisions for legal correctness, and that function becomes more likely to be filled by the Supreme Court. This would mean that if there is a positive relationship between circuit size and *unanimous* reversals, then we have evidence that circuit size negatively affects the ability of a circuit to maintain control of the legal quality of panel decisions.<sup>7</sup>

*Hypothesis 1a: The more judges a circuit has, the higher its unanimous reversal rate will be.*

The second component of circuit size is related to workload. Circuit judges frequently complain about the impact of workload, and much has been made of the impact of the increasing workload (Posner 1996) and reforms designed to cope with that increase while permitting the judges to focus on cases that present inter-

---

<sup>7</sup> A possibility that is difficult to explore directly is that a circuit may have judges who are more liberal or more conservative than all of the Supreme Court justices, and decisions by these judges could produce unanimous reversals. I attempt to control for this factor, albeit indirectly, by measuring a circuit’s ideological distance from the Supreme Court (a circuit with extreme conservatives or liberals would be of greater mean ideological distance) and by measuring a circuit’s ideological heterogeneity (circuits with more extremists should be more heterogeneous).

esting legal questions (Baker 1994). What is not clear is the independent effect workload may have on the frequency of reversal of judges in a circuit. Workload is often hypothesized to affect the behavior of appellate judges. Hettinger, Lindquist, and Martinek (2004) test the proposition that greater workload negatively affects the frequency of dissent of U.S. court of appeals judges, but find that it does not. Similarly, Posner (2000) finds no independent effect of workload on the number of summary reversals a circuit experiences, but this seems at least a little bit counterintuitive. Increased workload should increase the importance of efficient disposition of cases and attenuate the impact of ideological preferences on the decisions made by court of appeals judges. The busier court of appeals judges become, the more likely they are to seek the most efficient resolution of cases and the more emphasis they will place on accuracy and consistency as means toward the end of judicial efficiency. This would imply that the busier judges are, the *fewer* reversals they should experience.

*Hypothesis 2: The busier the circuit, the lower that circuit's reversal rate will be.*

The first set of hypotheses suggests that court of appeals judges attempt to make good law but can be affected by the environment in which they endeavor to do so. But court of appeals judges may not only be motivated by their desire to achieve legally correct results. They may prefer to act in accordance with their policy preferences. If this is the case, policy-based disagreements between court of appeals judges and the Supreme Court should produce more frequent reversals. The greater the ideological distance between the Supreme Court and the courts of appeals, the more likely the Supreme Court will reverse any given lower court decision. If this behavior is aggregated to the circuit level, then the greater the mean ideological distance between the Supreme Court and the circuit, the more often decisions from that circuit will be reversed.

*Hypothesis 3: The greater the ideological distance between the Supreme Court median and the circuit median, the more frequently a circuit will be reversed.*

If court of appeals judges and Supreme Court justices are motivated by their policy preferences, reversal of courts of appeals decisions will stem from ideological disagreement.

At the same time, the relationship between ideological distance and reversal frequency may be more nuanced than the straightforward relationship posited by Hypothesis 3. There may be a few judges or panels that lift the circuit's reversal rate independent of the composition of the rest of the circuit.

*Hypothesis 4: The more ideologically diverse a circuit is, the more reversals it will experience.*

If the en banc process modifies all decisions to the median of the circuit (or what the median of the circuit thinks the Supreme Court wants), then the heterogeneity of circuit preferences will not affect the reversal rate. If, however, there is no centripetal force, allowing each panel's decisions to stand pending review by the Supreme Court, then heterogeneity will matter because the circuit mean will not necessarily account for liberal judges in a conservative circuit or for conservative judges in a liberal circuit. If each panel's decision is final pending review by the Supreme Court, then circuit outliers will be able to make decisions that diverge from the circuit mean. This will happen more frequently in heterogeneous circuits and, once circuit ideology is controlled for (Hypothesis 3), would suggest that more heterogeneous circuits have more decisions reversed by the Supreme Court.

There is evidence for this behavior in the literature. Van Winkle (1996) argues that judges who are circuit outliers exploit opportunities that place them in panel majorities to behave contrary to the preference of the circuit median. If judges do this, and the circuit does not monitor such behavior using en banc review, then those decisions may be reversed by the Supreme Court if the Court also disagrees with the circuit outliers. Even if the behavior of circuit outliers is not so sophisticated, and they always behave sincerely, they will occasionally find themselves in panel majorities—more frequently if the circuit heterogeneity is greater.

## Data and Measures

The dependent variable in each analysis is the reversal rate of a circuit for each Supreme Court term between 1980 and 2002. This makes circuit-term the unit of analysis. The numerator, the number of reversals, was calculated using the Supreme Court database (Spaeth 2003). Supreme Court decisions that reversed the decisions of the circuit courts of appeals in the 1980–2002 terms were identified, and the number of times each circuit was reversed in a given term was counted. A relatively broad definition of reversal was used, including reversals in part and Supreme Court decisions that vacate lower court decisions.<sup>8</sup>

---

<sup>8</sup> The unit of analysis in the Supreme Court database is the case citation. To avoid including cases that were granted, vacated, and remanded (GVR) (Hellman 1984) by the Supreme Court, I excluded cases that were memorandum decisions that were vacated and remanded (there are very few GVRs left in the Supreme Court database once the unit of analysis is set to the case citation).

The denominator needed to calculate the reversal rate is the number of decisions that can be reversed. This was calculated using the number of merits terminations in a circuit, as provided by the *Federal Court Management Statistics* (Administrative Office of the United States Courts 1983, 1985, 1992, 1993, 1994, 1995, 1996, 1997, 1998, 1999, 2000, 2002). The reversal rate is the number of reversals a given circuit experienced in a given term divided by the number of merits terminations.<sup>9</sup> That number was multiplied by 100 to facilitate interpretation of the coefficients. Each circuit-term was then coded for circuit size, workload, ideological distance, and ideological heterogeneity.

Circuit size is not as clear a concept as one might expect. There is considerable variation across circuits in almost all components of size: population or geographical area, number of disputes that enter the federal court system, staff size, case volume, and number of judges. The primary objective was to measure the component(s) of the circuit that are most likely to affect its ability to create and maintain a coherent body of law. The size of the circuit affects the ability of judges to develop good working relationships with their colleagues. One of the problems created by circuit size is the inability of judges to work together and the lack of familiarity the judges have with their colleagues (Hellman 2003b; but see Kozinski 2003). This denies judges the ability to sit on panels with their colleagues and can impair the collegiality that can act to stabilize circuit results. The number of judgeships seemed to come closest to measuring the aspect of circuit size that affects the behavior of court of appeals judges (Posner 2000).

To assess the influence of workload on frequency of reversal, I used the number of merit terminations per active judge for each circuit. There are several different ways to attempt to measure how busy each judge in a circuit is, as Hellman (2003a) observes. The Administrative Office of the United States Courts, for example, relies on adjusted filings (taking into account reinstated cases and counting *pro se* filings as one-third of a filing) to determine how many judgeships each circuit should have. No measure perfectly captures how much work each active judge does on the circuit because the circuits vary considerably in their reliance on visiting and senior judges. Merit terminations per active judge will somewhat overstate the workload of judges in circuits that rely more heavily on visiting and senior judges, but any other measure will

---

<sup>9</sup> Merits terminations are measured over a statistical year (SY), which, since 1983, has run from October 1 to September 30. Generally speaking, cases decided by the courts of appeals during a statistical year would be candidates for reversal by the Supreme Court in the October Term (OT) of the same number. Cases decided during SY 1997, which started October 1, 1996, would be most likely to be reversed during OT 1997. The reversal rate was calculated as (# reversals in OT 19xx)/(# merits terminations in SY 19xx).

carry the same shortcoming. The advantage of using merits terminations per active judge is that it allows comparison to other quantitative work on the behavior of court of appeals judges (Hettinger et al. 2003, 2004; Posner 2000) and is a facially valid measure of the workload of the judges in a given circuit.

Ideological distance between the Supreme Court and each circuit proved to be extremely difficult to measure. Doing so requires a measure of the ideology of Supreme Court justices and a measure of the ideology of court of appeals judges, and placing those two measures on the same dimension so that the distance between the two can be measured. The Appendix explains in detail how this measure was derived. Briefly, I used inter-institution preference estimates developed by Bailey and Chang (2001) to place justices and appeals judges on the same dimension. Judges appointed under senatorial courtesy were assigned the ideology score of the home-state senator. Judges appointed in the absence of senatorial courtesy were assigned the ideology score of the appointing president.

Finally, to measure the heterogeneity of a circuit, I took the (sample) standard deviation of the circuit ideology. I expected the standard deviation of the judges' ideology measures to capture most of the aspects of what I hoped to measure; it has also been used to measure ideological heterogeneity in other work (Krehbiel 1991; Wahlbeck et al. 1998).<sup>10</sup>

Table 1 reports the summary statistics for judgeships, workload, ideology, and heterogeneity by circuit (1982–2002 for the Eleventh Circuit, 1980–2002 for all other circuits). There is a fair amount of variation across the circuits on all of the measures. The Seventh and Ninth Circuits are the circuits with the greatest average ideological distance, while the Fourth and Fifth Circuits are the circuits closest to the Supreme Court. This provides an important validity check on the measure of ideological distance. While the Ninth Circuit has the most judges, the Fifth and Eleventh are the busiest on a per-judge basis.

## Analysis

Panel data present a series of unique estimation challenges. Given that the data provide repeated observations,  $t = 1 \dots T$  for each individual,  $i = 1 \dots N$ , one must consider the possibility of

---

<sup>10</sup> One might suspect that, given that the denominator for standard deviation is  $N - 1$ , smaller circuits would have more ideological heterogeneity simply because of their size. More generally, one would expect a negative correlation between size and heterogeneity for this reason, but the correlation between number of judges and heterogeneity, while negative, is quite small ( $-0.0361$ ).

**Table 1.** Summary Statistics by Circuit, 1980–2002

	Number of Reversals	Reversal Rate <sup>a</sup>	Number of Unanimous Reversals	Number of Judgeships	Workload per Judge	Ideological Distance between Circuit and USSC <sup>b</sup>	Ideological Heterogeneity Within Circuit <sup>c</sup>
First Circuit	1.52 (1.47)	.28 (.33)	.52 (.79)	5.65 (.78)	301.04 (40.29)	1.80 (1.44)	1.78 (.55)
Second Circuit	3.96 (2.29)	.29 (.21)	1.61 (1.16)	12.65 (.78)	267.30 (48.45)	1.30 (.64)	1.24 (.17)
Third Circuit	3.48 (2.41)	.27 (.24)	1.39 (1.20)	12.78 (1.57)	334.35 (49.86)	1.54 (.93)	1.28 (.19)
Fourth Circuit	4.39 (2.25)	.28 (.24)	1.91 (1.35)	13.09 (2.25)	436.61 (154.23)	.95 (.70)	1.12 (.09)
Fifth Circuit	5.61 (2.90)	.22 (.16)	2.48 (1.65)	16.21 (1.13)	547.70 (155.07)	.86 (.31)	1.67 (.21)
Sixth Circuit	4.65 (2.12)	.24 (.13)	2.04 (1.61)	14.87 (1.87)	351.17 (51.66)	1.12 (.82)	1.67 (.21)
Seventh Circuit	3.70 (1.72)	.30 (.20)	1.78 (1.24)	10.65 (.78)	335.83 (53.27)	1.60 (1.07)	1.87 (.29)
Eighth Circuit	4.17 (2.41)	.34 (.28)	1.48 (1.38)	10.39 (.78)	385.00 (103.15)	1.18 (.87)	1.26 (.15)
Ninth Circuit	14.13 (5.59)	.48 (.32)	6.74 (3.71)	27.13 (1.94)	348.04 (115.16)	1.49 (.85)	1.55 (.16)
Tenth Circuit	2.56 (1.97)	.20 (.15)	1.30 (1.84)	10.78 (1.57)	325.96 (56.53)	1.08 (.71)	1.26 (.15)
Eleventh Circuit	3.76 (2.10)	.17 (.12)	1.71 (1.42)	12.00 (.00)	581.67 (200.54)	1.01 (.37)	1.09 (.23)
District of Columbia Circuit	4.17 (4.28)	.69 (.76)	1.30 (1.84)	11.83 (.39)	177.91 (39.11)	1.77 (.73)	1.89 (.33)
<b>All Circuits</b>	<b>4.69 (4.13)</b>	<b>.32 (.33)</b>	<b>2.03 (2.28)</b>	<b>13.18 (5.09)</b>	<b>346.47 (146.22)</b>	<b>1.31 (.88)</b>	<b>1.42 (.39)</b>

<sup>a</sup>Number of reversals per 100 cases terminated on the merits by the circuit in a given statistical year.

<sup>b</sup>Absolute distance between the Supreme Court median justice and the circuit median judge.

<sup>c</sup>Standard deviation of the ideology scores of the judges on the circuit (the values in the parentheses for this column are the standard deviation of the annual standard deviations).

Cell entries are averages across years for each circuit for the period 1980 and 2002 (1982–2002 for 11<sup>th</sup> Circuit). Standard deviations are shown in parentheses.

correlation within and across panels. In the context of the data analyzed here, one should be particularly sensitive to circuit-specific effects that may not be measured by the other independent variables. For example, the circuits vary considerably in the mix of populations they serve, which produces variations in the cases they decide that cannot be measured.

Two considerations in the presence of panel data are important. First, one must consider the treatment of the panel data itself. The primary classes of models for panel data are random- and fixed-effects models, and the choice between the two can have important implications for the results. Random-effects models are appropriate when one is interested in the relative importance of between- and within-group effects. More basically speaking, random-effects models are appropriate when one has a sample and is trying to generalize to the population. If, however, one is not seeking to generalize beyond the data collected, then fixed-effects models are an appropriate method of analysis. As the data analyzed here represent the population of circuits and not a sample, there is nothing to be gained in this context from using a random-effects model (see Zorn 2001). The primary advantage of a fixed-effects model is the ability to estimate an intercept for each circuit under analysis, conditioned on the independent variables. This is particularly useful in the context of ongoing debates about the Ninth Circuit. Much has been made about the Ninth Circuit's being reversed more frequently than the other circuits, but rarely has such an analysis focused on the frequency with which the Ninth is reversed *after* controlling for factors that should affect the reversal rate of any circuit. Allowing a circuit-specific estimate of the intercept (the baseline prediction for the number of reversals) can shed important light on the true status of the Ninth Circuit relative to its sister circuits. Once all of these factors are considered, it seems that a fixed-effects model is appropriate (Hsiao 2003).

The second problem is related to the first, but deals with the biases in estimates of the standard errors. Beck and Katz (1995) develop panel-corrected standard errors to account for the group-wise heteroskedasticity and contemporaneous correlation across panels in the errors (Beck & Katz 2004:4), but one must also be sensitive to the serial correlation of the errors. To address this possibility, the error structure is panel-specific AR (1), allowing the calculation of  $\rho_j$ , or an estimate of the correlation between  $e_t$  and  $e_{t-1}$  for each circuit.<sup>11</sup>

---

<sup>11</sup> As Beck and Katz (2004) note, the debate over the proper approach to time-series cross-sectional data is an ongoing one in the literature. Notably, they find the least-square dummy variable approach (the fixed-effects model used here) as a reasonably effective way to approach the issues created by time series cross-sectional data, and this approach has

**Table 2.** Predictors of Rate of Reversals, 1980–2002 Terms

Variable	$\beta$	Std. Error	$\beta$	Std. Error
Judgeships	.007*	.003	– 0.046***	.010
Merit Terminations per Judge (in 100s)	– 0.068***	.011	– 0.040**	.012
Ideological Distance	0.119***	.022	0.083***	.018
Ideological Dispersion	– 0.156***	.042	– 0.222***	.052
First Circuit			– 0.241**	.061
Second Circuit			– 0.006	.058
Third Circuit			– 0.015	.046
Fourth Circuit			0.077	.047
Fifth Circuit			0.184**	.067
Sixth Circuit			0.173*	.083
Seventh Circuit			0.056	.073
Ninth Circuit			0.924***	.178
Tenth Circuit			– 0.155*	.068
Eleventh Circuit			– 0.055	.055
District of Columbia Circuit			0.471	.299
Constant	0.536	.094	1.1485***	.126
R <sup>2</sup>	.338		.470	

Cell entries are OLS coefficients with panel-corrected standard errors in parentheses. \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ , two-tailed tests.  $N = 274$ .

## Results

If one pays no heed to circuit-specific (fixed) effects, the results are reflected in the first column of Table 2. If one considers a circuit's reversal rate to be the product of variables relating to circuit size, workload, and ideological composition, it becomes clear that all the factors are significant predictors of a circuit's reversal rate. The larger a circuit is, the higher its reversal rate. The busier the judges are on a circuit, the lower the reversal rate of that circuit. This supports the argument that judges in busier circuits are more likely than their colleagues in less busy circuits to eschew the pursuit of policy goals in favor of legal goals, as efficient disposition of cases becomes more important in busier circuits.

In addition, ideological distance is a strong predictor of a circuit's reversal rate. Circuits more ideologically distant from the Supreme Court are reversed at a greater rate than those that are closer to the Supreme Court. Additionally, and counter-intuitively, ideological dispersion is *negatively* related to a circuit's reversal rate. That is, controlling for ideological distance, the more diverse a circuit is ideologically, the *fewer* reversals it will experience.

Column 2 of Table 2 adds circuit-specific intercepts to ascertain whether the results observed in the first model are simply an artifact of circuit-specific factors. Generally speaking, the results observed in Column 1 of Table 2 persist once circuit-specific effects are included: circuits that are busier, are of greater ideological

computational advantages over others, particularly in unbalanced panels (seen here by fewer 11th Circuit observations than for the other circuits).

**Table 3.** Predictors of the Rate of Unanimous Reversals, 1980–2002 Terms

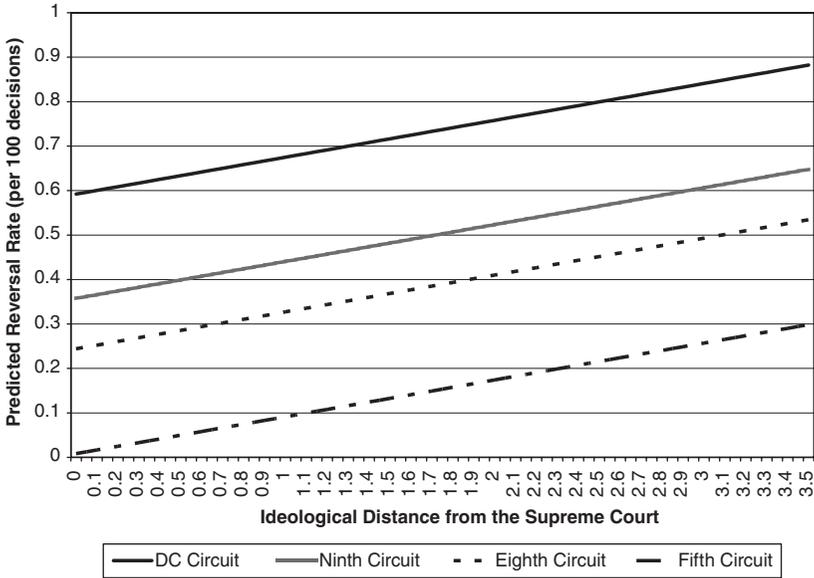
Variable	$\beta$	Std. Error	$\beta$	Std. Error
Judgeships	.004	.002	– 0.017	.007
Merit Terminations Per Judge (in 100s)	– .030***	.008	– 0.021*	.008
Ideological Distance	.025*	.012	0.011	.011
Ideological Dispersion	– .028	.025	– 0.074	.044
First Circuit			– 0.089	.052
Second Circuit			– 0.002	.036
Third Circuit			0.006	.035
Fourth Circuit			0.047	.032
Fifth Circuit			0.090	.057
Sixth Circuit			0.075	.054
Seventh Circuit			0.053	.048
Ninth Circuit			0.385**	.132
Tenth Circuit			– 0.034	.040
Eleventh Circuit			– 0.005	.046
District of Columbia Circuit			0.097	.087
Constant	.191	.069	0.459***	.097
R <sup>2</sup>	.145		.215	

Cell entries are OLS coefficients with panel-corrected standard errors in parentheses. \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ , two-tailed tests.  $N = 274$ .

distance, and are more ideologically homogeneous are reversed more frequently. Interestingly, once circuit-specific intercepts are included, a negative relationship between circuit size and reversal rate emerges. This is a particularly interesting finding because if one looks at the circuit intercepts, they demonstrate that larger circuits tend to experience higher reversal rates. If one holds the number of judges, workload, and ideological distance and dispersion at their overall means (not their circuit-specific means), the circuits with the highest predicted reversal rate are the Ninth, District of Columbia, Fifth, and Sixth Circuits.<sup>12</sup> The Ninth, Fifth, and Sixth Circuits are, respectively, the three largest circuits in the country. The results, then, appear inconclusive as to the relationship between circuit size and reversal rate.

One may remain dubious of the argument that the tested relationship between circuit size and the number of reversals is a fair test of the failure of judges in large circuits (or failure of the en banc proceedings in those circuits) to reach legally accurate decisions. A more rigorous test of this argument is to test the relationship between circuit size and the number of unanimous reversals. The dependent variable in these analyses is the number of unanimous reversals divided by the number of merits terminations in a given circuit (and multiplied by 100). Table 3 replicates the model used in Table 2, modeling first the independent variables without the fixed effects, and then adding the fixed effects. Looking at the model with circuit-specific intercepts (Column 2 of Table 3), the

<sup>12</sup> The Sixth Circuit is significant at  $p = 0.088$  (two tailed). The next-largest circuit, the Fourth, also has the next-highest reversal rate.



**Figure 1. Predicted Reversal Rate by Ideological Distance from Supreme Court: D.C., Eighth, Ninth, and Fifth Circuits**

predictors of the rate of unanimous reversals are similar to those of a circuit’s overall reversal rate, with two notable exceptions. Circuits farther away on the ideological spectrum are not unanimously reversed more frequently (the effect is significant without the circuit-specific intercepts), and circuits that are more ideologically homogeneous are unanimously reversed no more often than ideologically heterogeneous circuits. Similar to the results for all reversals, busier circuits are unanimously reversed at a lower rate. The same pattern for circuit size emerged for unanimous reversals that was found for all reversals. These results suggest that unanimous reversals are not a distinct class of Supreme Court decisions, contradicting other explanations for the frequency of unanimous reversals. Unanimous reversals may be nothing more than reversals of decisions that are either more liberal or more conservative than any justice of the Supreme Court is willing to accept.

To fully assess the impact of the different variables on a circuit’s reversal rate, it may be useful to look at the predicted reversal rate for different circuits. This may prove particularly useful when attempting to assess the competing influences of the circuit-specific factors and the overall variables. If the independent variables are set at their circuit-specific means, one can see the impact of the variation in the number of judges and its impact on the individual circuit intercepts. Figure 1 illustrates the result of increases in ideological distance on reversal rate. The range of ideological

distance is 0.01 to 3.5. If one were to increase the ideological distance of the Eighth Circuit from the minimum to the maximum ideological distance, its reversal rate per 100 cases increases from 0.24 to 0.53 (from 0.24 to 0.53% of the cases being reversed) While the difference may seem trivial in absolute terms, the impact of the move increases the reversal rate nearly three-fold. The same increase on the Ninth Circuit would drive the reversal rate per 100 cases from 0.35 to 0.64, nearly doubling it.

## Discussion

Overall, there is mixed support for the argument that judges attempt to make good law and that their attempts to do so are either enhanced or impeded by the organizational features of their circuits. The evidence that the size of a circuit is related to its reversal rate is difficult to evaluate. While there is a negative relationship between the number of judges and a circuit's reversal rate, the circuit-specific effects suggest that larger circuits (the Fourth, Fifth, Sixth, and Ninth Circuits) are reversed at a greater rate than their sister circuits. On the other hand, there is consistent evidence that the workload of each judge is negatively related to the reversal rate of a circuit (though not to the frequency of unanimous reversals). This can be read as evidence that busier judges place greater value on attempting to reach legally correct decisions and eschew behavior according to their policy preferences, behavior that may generate more reversals. This is an interesting finding relative to the extant literature on the subject. Hettinger, Lindquist, and Martinek (2004), for example, find no relation between workload and the decision to dissent from a panel decision. The results presented here, though, suggest that workload pressures have an effect on the behavior of court of appeals judges.

Turning to assessment of evidence for the hypotheses related to ideological disagreement as the source of a circuit's reversal rate, it is clear that circuit ideological distance and circuit heterogeneity are predictors of a circuit's reversal rate. This suggests that reversals are governed primarily by ideological disagreement between the Supreme Court and the courts of appeals. Reversal rate is also affected by the internal homogeneity of a circuit. This finding is somewhat unexpected. The literature suggests that the more diverse a circuit is, at least in terms of number of judges who differ from the circuit mean, the more diverse the output of the circuit will be (if one aggregates the panel decisions). This would suggest that more diverse circuits, once the ideology of the circuit is controlled for, would produce more decisions the Supreme Court would seek to reverse. But a negative relationship between

ideological diversity and number of reversals is more difficult to explain. It may be the case that judges in more diverse circuits feel greater pressure to conform to circuit norms, lest the output of the circuit be too inconsistent for litigants, attorneys, and judges to follow, but this is mere conjecture.

A more plausible explanation for this phenomenon may come from the Supreme Court's method of auditing decisions from the different circuits. Much of the work to date on the relationship between the Supreme Court and the courts of appeals has assumed that the Supreme Court observes the ideology of the opinion author or the panel ideology as a whole (Cameron et al. 2000). But the Supreme Court may pay attention to the signal sent by the entire circuit (Haire et al. 2001), more closely auditing circuits more ideologically distant than those closer (as the results presented here suggest). The ideological heterogeneity of a circuit may affect the clarity of such a signal: the Supreme Court may find it easier to interpret a decision from an ideologically homogeneous circuit than from one with considerable heterogeneity. To Krehbiel (1991), heterogeneity of an agent (for him, committees of Congress) is a virtue. If members of a heterogeneous committee are able to reach consensus on legislation, the rest of the House can assume that the legislation reflects the members' preferences and is therefore more likely to be considered under a closed rule. In the same way, if liberal and conservative judges agree on an outcome (panels mixing ideologies would be more likely in heterogeneous circuits), then the Supreme Court can trust that the outcome is an acceptable one. If Representatives Ron Dellums (D-CA) and Edward Hebert (D-LA), both of the House Armed Services Committee, could agree on legislation in the 1970s, then the House could conclude that the legislation represented the interests of the House as a whole. By the same logic, if Judges Stephen Reinhardt and Alex Kozinski (a noted liberal and conservative, respectively, on the Ninth Circuit) can agree on the disposition of a case, the Supreme Court may have little reason to doubt that the outcome of the case was correct.

Much has been made of the costs the Supreme Court incurs by deciding to review lower court decisions (Cameron et al. 2000; Spitzer & Talley 2000), so there may be smaller costs associated with reviewing (and reversing) circuits that generate clearer signals, as indicated by a more homogeneous ideology. This would mean that, all else being equal, the costs of reviewing the decisions of heterogeneous circuits would be higher than the costs associated with reviewing the decisions of more homogeneous circuits, which would lead to more frequent reversals of ideologically uniform circuits (again, having controlled for circuit ideological distance from the Supreme Court). This would also explain how circuit

outliers can exploit their status as temporary panel majorities without fear of reversal by an en banc review (Van Winkle 1996), and such behavior may benefit the circuit as a whole by raising the costs associated with Supreme Court review.

This finding may also have implications for understanding how en banc review, or the threat of such a review, functions in different circuits. Previous work has speculated that one of the functions of en banc review is to correct erroneous panel decisions (George 1999), but it is difficult to square this proposition with the finding that the Supreme Court is more likely to review cases decided en banc than those decided by panels (George & Solimine 2001). This may imply that en banc review occurs in more interesting cases—the kind of cases the Supreme Court is also more likely to hear—and this effect may trump the efforts by the courts of appeals to correct erroneous decisions. The results here may suggest that en banc review serves different purposes in both homogeneous and heterogeneous circuits. In heterogeneous circuits, it may serve to correct errors made by panels that diverge from the circuit mean, reducing reversal rates of heterogeneous circuits. In homogeneous circuits, such error correction will be needed less often, so it may serve to identify important legal issues in those circuits, increasing reversal rates of those circuits. This may help explain the inverse relationship between a circuit's heterogeneity and its reversal rate, but the subject certainly warrants further investigation.

Finally, these results may help shed light on a continuing controversy—the source of the Ninth Circuit's high number of reversals by the Supreme Court. The results in Tables 2 and 3 demonstrate that the Ninth Circuit stands out from most of the other circuits. The choice of the excluded category (here, the Eighth) is arbitrary, but it has no effect on estimation of the other coefficients. If one changes the excluded category to the Fifth Circuit, the circuit with the next-largest intercept, the effect of the Ninth Circuit is still positive and significant.<sup>13</sup> There is substantial evidence, then, that the rate at which the Ninth Circuit is reversed exceeds what one would expect beyond controls for circuit size, workload, ideological distance, and ideological heterogeneity. The source of its high reversal rate, then, transcends these characteristics, which may suggest that further exploration of this phenomenon is still warranted.

---

<sup>13</sup> The circuit-specific intercept is the reversal rate expected when all the independent variables are set to zero. As discussed earlier, once one sets the independent variables to the circuit-specific means, the Fifth is one of the circuits with the lowest expected reversal rate, primarily due to its very high workload per judge.

## Conclusion

Analysis of the variation in reversal rates of the U.S. courts of appeals offers leverage for understanding the dynamics of judicial hierarchies more generally. Most intermediate appellate courts hear cases in subsets of the entire bench: the Australian Federal Court assembles in ad hoc panels of three Federal Court judges. In Canada and England, the intermediate appellate courts have full-time judges, but they also serve in panels. If there is variation in the rate at which judges, panels, or entire courts are reversed, then the sources of that variation offer the opportunity to shed light on what motivates the judges on intermediate courts of appeal and the degree to which both law and policy motivate them. In addition, cross-national studies would offer the opportunity to understand the role of influences that might be constant within a country, including the desire for promotion to the highest court. If one assumes that all (or almost all) lower court judges desire higher office, then evaluating that influence on their behavior is difficult unless one looks across systems, where opportunities for promotion vary substantially.<sup>14</sup> The model laid out here is an initial attempt to lay the groundwork for future research into what motivates appellate judges. Like high court judges, they are motivated by their desire to make policy. But lower court judges also appear to attempt to make clear and consistent law, and their attempts to do so are affected by the court on which they serve. These findings suggest that the determinants of behavior of intermediate appellate court judges are quite complex, and scholars need to think carefully about simply transplanting models of judicial behavior from higher to lower court judges.

As our understanding of the common features of different political systems evolves, political scientists have come to appreciate the importance of the judiciary in the development of policy. Particularly when courts engage in constitutional interpretation, they have the power to make policy that cannot be overturned by normal legislative means. To some, this raises the prospect of a government of judges (Stone Sweet 2000; Volcansek 2000). To others, such a concern is premature given that the ideal points of the judges (the point in space where a political actor would most prefer policy be located) are likely located in the “unanimity core of the other veto players” (Tsebelis 2002:227). But to date, this literature has focused almost exclusively on the behavior of judges on high courts, and on constitutional courts in particular. In many systems, however, the lower courts shape the contours of the terrain on

---

<sup>14</sup> Particularly in systems where judges and justices are forced to retire at a certain age, lower court judges may see promotion as an attainable goal.

which high courts make their impact on policy. Atkins (1993) suggests that the English Court of Appeal is quite active in this regard, and research on the certiorari process in the U.S. Supreme Court suggests that the justices attend to signals (particularly dissenting opinions) provided by the U.S. courts of appeals (Caldeira et al. 1999). But intermediate appellate courts are important in their own right as well. As lower courts are the final arbiter of the vast majority of disputes, their collective impact may rival that of the court that sits above them.

Much of the reason that lower appellate courts have been understudied is because their decisionmaking patterns are assumed rather than tested. I demonstrate here that the factors that motivate intermediate appellate judges are more nuanced than traditionally believed. Judges who sit directly below their nation's highest court are clearly motivated by their policy preferences, but there is evidence to suggest that they also try to make good law, an effort that is affected by their work environment. Understanding what motivates intermediate appellate judges offers the opportunity to better appreciate courts not just as players in national politics, but as a system of vertically organized units staffed by professionals whose motivations are affected by their location within the hierarchy. That is, not all judges behave as members of the high courts do, and appreciating the patterns of the behavior of lower court judges enhances our understanding of the impact of institutional structure on the preferences and behaviors of political actors.

### **Appendix: Measuring Ideological Distance between the U.S. Supreme Court and the U.S. Courts of Appeals**

Measuring the ideological distance between the Supreme Court and the courts of appeals required three components: measuring ideology for each level, and checking that those two measures are on the same dimension. Turning first to the measure of the ideology of court of appeals judges, I used the measure of court of appeals ideology developed by Giles, Hettinger, and Peppers (2001) (hereafter GHP). The GHP approach assigns court of appeals judges scores related to the circumstances of their appointment. Judges appointed in the absence of senatorial courtesy, including judges on the D.C. Circuit Court of Appeals, take on the ideology score of the appointing president (for GHP, the NOMINATE score). Judges appointed in the presence of senatorial courtesy take on the ideal point of the senator (if there are two of the president's party, the average score) responsible for their appointment. GHP find that such a method of preference estimation

better predicts the votes of judges than one that solely accounts for the president's policy preferences. One might argue that there has been an evolution in how presidents select their court of appeals judges, with Ronald Reagan, in particular, focusing on policy rather than partisan criteria (Goldman 1997), but GHP contend that presidents who preceded Reagan "talked less about policy in the selection process . . . but in the end they selected federal judges who reflected their policy preferences whenever possible" (2001:638). Their evidence builds a compelling empirical and theoretical case for using this as a measure of ideology for court of appeals judges. Such an approach has a decided advantage over approaches that rely only on the appointing president or those that build scores out of a judge's attributes, including work history, region, and religion (Cameron et al. 2000).

This is an extremely attractive option for court of appeals judges, but there is no way to compare GHP scores to ideology scores for the Supreme Court, which is necessary to measure ideological distance between the two levels. Two major methodological advances have occurred in Supreme Court ideal point estimation over the past several years. First, there has been a proliferation of vote-based measures of ideology for Supreme Court justices to replace Segal-Cover (Segal & Cover 1989) scores. These measures (Bailey & Chang 2001; Martin & Quinn 2002) have the important advantage of varying over a justice's career, allowing for changes in justices' ideology, something previously difficult to measure.<sup>15</sup>

The second major advance closely accompanies the first, and that is the development of inter-institutional preference estimation. As Segal (1997) noted, comparing measures of ideology across institutions is notoriously difficult. But Bailey and Chang, using positions taken by the president on legislation before the Senate and amicus briefs by the Solicitor General before the Supreme Court, have developed ideal point estimates for the president, Senate, and Supreme Court that lie on the same dimension (Bailey & Chang 2001). This is an important advancement because it allows the measurement of Supreme Court justices on the same dimension as that of court of appeals judges (if one measures court of appeals judges by the president and senator(s) responsible for their selection) and permits the creation of a measure of ideological distance between the Supreme Court and the judges of the courts of appeals.

I adapted the GHP methodology by using Bailey and Chang (2001) ideal point estimates for presidents and senators (GHP use

---

<sup>15</sup> Segal-Cover scores use editorials surrounding justice confirmation processes to gauge the justices' preferences. Vote-based measures rely on the votes of the justices on the Court to place the justices on a liberal-conservative continuum.

NOMINATE scores, but their approach is generalizable). The circuit ideology is the ideology of the median active judge on the circuit in a given year. The Supreme Court ideology is the Bailey and Chang score for the median justice. The ideological distance between the Supreme Court and a given circuit  $i$  in year  $t$  is the absolute value of the distance between the two.

I thank Michael Bailey for his advice on using the Supreme Court scores. To address the possibility of a lag between the time cases are decided by the courts of appeals and the time they are decided by the Supreme Court, court of appeals ideology is measured by calendar year, while Supreme Court ideology is measured by term. One would expect a lag of about a year between the two decisions, and that is what this assumes. Because Bailey-Chang scores are only calculated through 1996 (the 1995 term), I fixed the Supreme Court's ideology at that level (0.03) for 1996–2002.

## References

- Abe, Masaki (1995) "The Internal Control of a Bureaucratic Judiciary: The Case of Japan," 23 *International J. of the Sociology of Law* 303–20.
- Administrative Office of the United States Courts (1983, 1985, 1992, 1993, 1994, 1995, 1996, 1997, 1998, 1999, 2000, 2002) *Federal Court Management Statistics*. Washington, D.C.: U.S. Government Printing Office.
- Atkins, Burton (1990) "Interventions and Power in Judicial Hierarchies: Appellate Courts in England and the United States," 24 *Law & Society Rev.* 71–103.
- (1993) "Alternative Models of Appeal Mobilization in Judicial Hierarchies," 37 *American J. of Political Science* 780–98.
- Bailey, Michael, & Kelly Chang (2001) "Comparing Presidents, Senators, and Justices: Interinstitutional Preference Estimation," 17 *J. of Law, Economics, and Organization* 477–516.
- Baker, Thomas E. (1994) *Rationing Justice on Appeal: Problems of the US Court of Appeals*. St. Paul, MN: West Publishing.
- Baum, Lawrence (1997) *The Puzzle of Judicial Behavior*. Ann Arbor: Univ. of Michigan Press.
- Beck, Nathaniel, & Jonathan N. Katz (1995) "What to Do (and Not to Do) with Time-Series Cross-Section Data," 89 *American Political Science Rev.* 634–47.
- Beck, Nathaniel (2004) "Time-Series Cross-Section Issues: Dynamics, 2004." Presented at the Annual Meeting of the Political Methodology Section of the American Political Science Association, Palo Alto, CA.
- Caldeira, Gregory A., et al. (1999) "Sophisticated Voting and Gate-Keeping in the Supreme Court," 15 *J. of Law, Economics, and Organization* 549–72.
- Cameron, Charles, et al. (2000) "Strategic Auditing in a Political Hierarchy: An Informational Model of the Supreme Court's Certiorari Decisions," 43 *American Political Science Rev.* 162–85.
- Cross, Frank B. (2003) "Decisionmaking in the U.S. Circuit Courts of Appeals," 91 *California Law Rev.* 1457–515.
- Cross, Frank B., & Emerson H. Tiller (1998) "Judicial Partisanship and Obedience to Legal Doctrine: Whistleblowing on the Federal Courts of Appeals," 107 *Yale Law J.* 2155–76.

- Epp, Charles R. (1998) *The Rights Revolution: Lawyers, Activists and Supreme Courts in Comparative Perspective*. Chicago: Univ. of Chicago Press.
- Epstein, Lee, et al. (2003) "The Norm of Prior Judicial Experience and Its Consequences for Career Diversity on the U.S. Supreme Court," 91 *California Law Rev.* 903–66.
- Fleming, Roy B. (2004) *Tournament of Appeals: Granting Judicial Review in Canada*. Vancouver: UBC Press.
- Fleming, Roy B., & Glen S. Krutz (2002) "Selecting Appeals for Judicial Review in Canada: A Replication and Multivariate Test of American Hypotheses," 64 *J. of Politics* 232–48.
- George, Tracey E. (1999) "The Dynamics and Determinants of the Decision to Grant En Banc Review," 74 *Washington Law Rev.* 213–74.
- George, Tracey E., & Michael E. Solimine (2001) "Supreme Court Monitoring of the United States Court of Appeals En Banc," 9 *Supreme Court Economic Rev.* 171–204.
- Giles, Micheal W., et al. (2001) "Picking Federal Judges: A Note on Policy and Partisan Selection Agendas," 54 *Political Research Q.* 623–41.
- Ginsburg, Tom (2003) *Judicial Review in New Democracies: Constitutional Courts in Asian Cases*. New York: Cambridge Univ. Press.
- Goldman, Sheldon (1966) "Voting Behavior on the United States Courts of Appeals, 1961–1964," 60 *American Political Science Rev.* 374–83.
- (1975) "Voting Behavior on the United States Court of Appeals Revisited," 69 *American Political Science Rev.* 491–506.
- (1997) *Picking Federal Judges: Lower Court Selection from Roosevelt through Reagan*. New Haven: Yale Univ. Press.
- Greene, Ian, et al. (1998) *Final Appeal: Decision-Making in Canadian Courts of Appeal*. Toronto: James Lorimer.
- Guarnieri, Carlo, & Patrizia Pederzoli (2002) *The Power of Judges: A Comparative Study of Courts and Democracy*, Ed. in English C. A. Thomas. Oxford: Oxford Univ. Press.
- Haire, Susan, et al. (2001) "Hierarchical Relationships in the Federal Courts." Presented at the Annual Meeting of the American Political Science Association, San Francisco, CA.
- Hellman, Arthur D. (1984) "Granted, Vacated, and Remanded: Shedding Light on a Dark Corner of Supreme Court Practice," 67 *Judicature* 389–401.
- (2000) "Getting It Right: Panel Error and the En Banc Process in the Ninth Circuit Court of Appeals," 34 *University of California, Davis Law Rev.* 425–69.
- (2003a) "Assessing Judgeship Needs in the Federal Courts of Appeals: Policy Choices and Process Concerns," 5 *J. of Appellate Practice and Process* 239–70.
- (2003b) Statement in Hearing on HR 2723, Ninth Circuit Court of Appeals, Judgeship and Reorganization Act of 2003. 108th Cong., 1st sess., 21 October. <http://www.house.gov/judiciary/90125.PDF> (accessed 27 March 2004).
- Herald, Marybeth (1998) "Reversed, Vacated and Split: The Supreme Court, the Ninth Circuit, and Congress," 77 *Oregon Law Rev.* 405–96.
- Hettinger, Virginia A., et al. (2003) "Acclimation Effects and Separate Opinion Writing in the U.S. Courts of Appeals," 84 *Social Science Q.* 792–810.
- (2004) "Comparing Attitudinal and Strategic Accounts of Dissenting Behavior on the U.S. Courts of Appeals," 48 *American J. of Political Science* 123–37.
- Hirschl, Ran (2004) *Towards Juristocracy: The Origins and Consequences of the New Constitutionalism*. Cambridge: Harvard Univ. Press.
- Hsiao, Cheng (2003) *Analysis of Panel Data*, 2nd ed. Cambridge: Cambridge Univ. Press.
- Klein, David E. (2002) *Making Law in the United States Courts of Appeals*. New York: Cambridge Univ. Press.
- Klein, David E., & Robert J. Hume (2003) "Fear of Reversal as an Explanation of Lower Court Compliance," 37 *Law & Society Rev.* 579–606.

- Kornhauser, Lewis A. (1995) "Adjudication by a Resource-Constrained Team: Hierarchy and Precedent in a Judicial System," 65 *Southern California Law Rev.* 1605–28.
- Kozinski, Alex (2003) Statement in Hearing on HR 2723, Ninth Circuit Court of Appeals Judgeship and Reorganization Act of 2003. 108th Cong., 1st sess., 21 October. <http://www.house.gov/judiciary/90125.PDF> (accessed 27 March 2004).
- Krehbiel, Keith (1991) *Information and Legislative Organization*. Ann Arbor: Univ. of Michigan Press.
- Martin, Andrew D., & Kevin M. Quinn (2002) "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953–1999," 10 *Political Analysis* 134–53.
- McCormick, Peter (1992) "The Supervisory Role of the Supreme Court of Canada: Analysis of Appeals from Provincial Courts of Appeal, 1949–1990," 3 *Supreme Court Law Rev.* 1–28.
- (2000) *Supreme at Last: The Evolution of the Supreme Court of Canada*. Toronto: James Lorimer.
- Mikulski, Barbara (2003) "Senator Barbara Mikulski Opposes Nomination of Virginia Judge to Maryland Seat on 4th Circuit Court of Appeals." <http://mikulski.senate.gov/record.cfm?id=214117> (accessed 15 September 2004).
- Perry, H. W. Jr. (1991) *Deciding to Decide: Agenda Setting in the United States Supreme Court*. Cambridge: Harvard Univ. Press.
- Pierce, Jason (2002) "Judicial Entrepreneurism and the Politics of Institutional Change: An Analysis of the Recent Judicial Role Transformation in the High Court of Australia." Ph.D. dissertation, University of Texas.
- Posner, Richard A. (1996) *The Federal Courts: Challenge and Reform*. Cambridge: Harvard Univ. Press.
- (2000) "Is the Ninth Circuit Too Large? A Statistical Study of Judicial Quality," 29 *J. of Legal Studies* 711–9.
- Provine, Doris Marie (1980) *Case Selection in the United States Supreme Court*. Chicago: Univ. of Chicago Press.
- Salzberger, Eli, & Paul Fenn (1999) "Judicial Independence: Some Evidence from the English Court of Appeal," 42 *J. of Law and Economics* 831–47.
- Scalia, Antonin (1998) Submitted Testimony to the Commission on Structural Alternatives for the Federal Courts of Appeals. <http://www.library.unt.edu/gpo/csafca/hearings/submitted/pdf/Scalia1.pdf> (accessed 25 July 2004).
- Segal, Jeffrey A. (1997) "Separation-of-Powers Games in the Positive Theory of Congress and the Courts," 91 *American Political Science Rev.* 28–44.
- Segal, Jeffrey A., & Albert D. Cover (1989) "Ideological Values and the Votes of US Supreme Court Justices," 83 *American Political Science Rev.* 557–64.
- Segal, Jeffrey A., & Harold J. Spaeth (2002) *The Supreme Court and the Attitudinal Model Revisited*. New York: Cambridge Univ. Press.
- Shapiro, Martin, & Alec Stone Sweet (2002) *On Law, Politics and Judicialization*. Oxford: Oxford Univ. Press.
- Songer, Donald R., et al. (1994) "The Hierarchy of Justice: Testing a Principal-Agent Model of Supreme Court-Circuit Court Interactions," 36 *American J. of Political Science* 673–96.
- Spaeth, Harold J. (2003) "Original United States Supreme Court Database (1953–2002 Terms)." <http://www.as.uky.edu/polisci/ulmerproject/sctdata.htm> (accessed 13 October 2005).
- Spitzer, Matt, & Eric Talley (2000) "Judicial Auditing," 29 *J. of Legal Studies* 649–83.
- Stone Sweet, Alec (2000) *Governing with Judges: Constitutional Politics in Europe*. Oxford: Oxford Univ. Press.
- Neal, Tate C., & Torbjörn Vallinder, eds. (1995) *The Global Expansion of Judicial Power*. New York: New York Univ. Press.

- 
- Tsebelis, George (2002) *Veto Players: How Political Institutions Work*. New York: Russell Sage Foundation.
- Van Winkle, Steven R. (1996) "Three-Judge Panels and Strategic Behavior on the US Courts of Appeals." Presented at the Conference on the Scientific Study of Judicial Politics, St. Louis, MO.
- Volcansek, Mary L. (2000) *Constitutional Politics in Italy: The Constitutional Court*. New York: St. Martin's Press.
- Wahlbeck, Paul J., et al. (1998) "Marshalling the Court: Bargaining and Accommodation on the United States Supreme Court," 42 *American J. of Political Science* 294–315.
- Wasby, Stephen L. (1998) "The Ninth Circuit and the Supreme Court: Relations Between Higher and Lower Courts." Paper presented at the Annual Meeting of the American Political Science Association, Boston, MA.
- (2001) "The Supreme Court and Courts of Appeals En Bancs," 33 *McGeorge Law Rev.* 17–73.
- Zorn, Christopher J. W. (2001) "Generalized Estimating Equation Models for Correlated Data: A Review with Applications," 45 *American J. of Political Science* 470–90.

**Kevin Scott** is an Assistant Professor in the Department of Political Science at Texas Tech University. His research focuses on the U.S. courts of appeals and the relationship between the federal courts and Congress. His recent work appears in the *Virginia Law Review* and in *The Rehnquist Legacy* (ed. Craig Bradley, Cambridge University Press, forthcoming).