JUDGED BY THE COMPANY YOU KEEP: AN EMPIRICAL STUDY OF THE IDEOLOGIES OF JUDGES ON THE UNITED STATES COURTS OF APPEALS

Corey Rayburn Yung*

Abstract: Although there has been an explosion of empirical legal scholarship about the federal judiciary, with a particular focus on judicial ideology, the question remains: how do we know what the ideology of a judge actually is? For federal courts below the U.S. Supreme Court, legal academics and political scientists have offered only crude proxies to identify the ideologies of judges. This Article attempts to cure this deficiency in empirical research about the federal courts by introducing a new technique for measuring the ideology of judges based upon judicial behavior in the U.S. courts of appeals. This study measures ideology, not by subjectively coding the ideological direction of case outcomes, but by determining the degree to which federal appellate judges agree and disagree with their liberal and conservative colleagues at both the appellate and district court levels. Further, through regression analysis, several important findings related to the Ideology Scores emerge. First, the Ideology Scores in this Article offer substantial improvements in predicting civil rights case outcomes over the leading measures of ideology. Second, there were very different levels and heterogeneity of ideology among the judges on the studied circuits. Third, the data did not support the conventional wisdom that Presidents Ronald Reagan and George W. Bush appointed uniquely ideological judges. Fourth, in general judges appointed by Republican presidents were more ideological than those appointed by Democratic presidents. Fifth, attendance at a higher-ranked law school was strongly correlated with liberalism on the bench for appointments of both Republican and Democratic presidents. Sixth, prior work experience in the government (outside the judiciary) indicated liberal judicial voting.

INTRODUCTION

The key front in recent judicial confirmation battles in the Senate has been over whether a nominee's ideology was "radical" or "extreme."[1] On the eve of President Barack Obama's announcement of his nomination to replace Justice David Souter on the U.S. Supreme Court, conservative groups branded potential candidates, such as now-Justices Sonia Sotomayor and Elena Kagan, and Judge Diane Wood, as having views "disturbingly out of the mainstream."[2] Similar charges were made against Chief Justice John Roberts and Justice Samuel Alito before they were confirmed by the Senate.[3] After the recent retirement of Justice John Paul Stevens, threats of filibusters were made against several of the rumored nominees due to questions about ideology.[4] For appointments to other federal courts, challenges to a nominee's ideology have been at the heart of political confrontations that have slowed the pace of confirmations to a crawl.[5] Illustrating the absurdity of the political theater surrounding the confirmation process, critics contended that Professor Goodwin Liu, who was nominated to the Ninth Circuit, was out of the political mainstream because Liu had argued that Justice Alito's ideology was "at the margin of the judicial spectrum, not the mainstream."[6] Although ideology has become the focal point of discussions of the federal judiciary, the question remains as to whether it can be said with any confidence what a judge's ideology actually is. For federal courts below the U.S. Supreme Court, legal academics and political scientists have offered only crude proxies to identify the ideologies of judges.[7]

---

[1] *See* Editorial, *Battle for Courts Far From Over*, CHATTANOOGA FREE TIMES, June 11, 2005, at B7.

[2] Dana Bash, *Conservatives Prepare for Supreme Battle*, CNN (May 1, 2009), http://www.cnn.com/2009/POLITICS/05/01/gop.scotus/index.html.

[3] *See* Jeffrey Lord, *Conservatives and the Court: Their Moment Is at Hand*, PHILA. INQUIRER, Nov. 8, 2005, at A11 (discussing Justice Alito's nomination to the Supreme Court); Maeve Reston, *Hearings Likely to Be Dramatic*, PITTSBURGH POST-GAZETTE, Sept. 4, 2005, at A10 (describing the confirmation hearings of Justice Roberts).

[4] Laura Litvan, *GOP Readies for Fight over Stevens' Successor*, BOS. GLOBE (Apr. 12, 2010), http://www.boston.com/news/nation/washington/articles/2010/04/12/gop_readies_for_fight_over_stevens_successor/?s_campaign=8315.

[5] *See* Editorial, *Partisan Payback*, L.A. TIMES, Nov. 17, 2009, at A34.

[6] Robert Barnes, *Law Professor Goodwin Liu May Be a Test Case for Obama Judicial Picks*, WASH. POST (Mar. 22, 2010), http://www.washingtonpost.com/wp-dyn/content/article/2010/03/21/AR2010032102581.html; *see* Carol J. Williams, *A Big Legal Resume, Getting Bigger; Goodwin Liu, 39, May Soon Be the Youngest Judge on 9th Circuit*, L.A. TIMES, Mar. 9, 2010, at A1.

[7] *See* Joshua B. Fischman & David S. Law, *What Is Judicial Ideology, and How Should We Measure It?*, 29 WASH. U. J.L. & POL'Y 133, 170 (2009) ("[I]t is obvious that party of appointment and other proxies can be rather crude measures of ideology.").

The most commonly-used measure of the ideology of federal judges is the political party of the appointing president.[8] All appointed judges are scored by a binary measure that transfers the political party of the appointing president to the judge.[9] Notably, it identifies recent and current Justices Clarence Thomas, John Paul Stevens, Sandra Day O'Connor, Antonin Scalia, Samuel Alito, Harry Blackmun, and David Souter as ideological equals because they were appointed by Republican presidents.[10] Common Space Scores[11] have been hailed as a substantial improvement over the proxy measure of the appointing president because they incorporate the ideologies of the nominee's home-state senators with the president's ideology.[12] Yet recent research has called into question whether those scores offer any significant measurable gains.[13]

Justice Sotomayor's history illustrates one of the oddities of the two major proxy measures. When she was nominated by President George H.W. Bush to a position on a federal district court, both of the leading measures would have labeled her a conservative.[14] With subsequent appointments to a federal appellate court and the Supreme Court, however, she is now considered a liberal.[15] That seeming transformation

---

[8] *See id.* at 167–68 ("The most popular proxy for a judge's ideology, however, has been the party of the official who appointed the judge . . . . The appointing-party measure has been especially dominant in studies of the federal courts."); Daniel R. Pinello, *Linking Party to Judicial Ideology in American Courts: A Meta-Analysis*, 20 Just. Sys. J. 219, 219 (1999) (quoting Gregory N. Flemming et al., *An Integrated Model of Privacy Decision Making in State Supreme Courts*, 26 Am. Pol. Q. 35, 40 (1998)) ("[T]he most common approach . . . has been to use the party affiliation of a judge or of the governor who appointed the judge.").

[9] *See* Frank B. Cross, Decision Making in the U.S. Courts of Appeals 19 (2007) ("[R]esearchers have presumed that judges appointed by Democrats are ideologically liberal whereas those appointed by Republicans are ideologically conservative.").

[10] *See id.* Each of these seven Justices was appointed by a Republican president, meaning that they were coded as having the same ideology score under the appointing president proxy. *See id.*

[11] *See* Michael W. Giles et al., *Picking Federal Judges: A Note on Policy and Partisan Selection Agendas*, 54 Pol. Res. Q. 623, 631 (2001) (calculating a Common Space Score "for the state congressional delegation of the President's party in the year of the judge's appointment").

[12] *See id.* (labeling Common Space Scores as the "best currently available measure for circuit court judicial ideology").

[13] *See* Fischman & Law, *supra* note 7, at 200–01 ("Although there has been much debate about the relative merits of common space scores as opposed to party of appointment as measures of judicial ideology[,] . . . performance differences between these two proxy measures are quite small in practice. Common space scores provide a better fit to the asylum voting data than party of appointment, but the difference is slight. Both measures yield almost identical estimates of the impact of ideology on how judges vote.").

[14] *See* Cross, supra note 9, at 19.

[15] *See id.*

was not based upon any of her decisions—it was determined solely by the politics of President Clinton, President Obama, and the U.S. senators from New York.[16] Had she remained a district court judge, she would forever be considered a conservative in academic study. Therefore, the political party of the appointing president and Common Space Scores are essentially useless proxy measures in situations such as confirmation battles because the measures gauge judges with the same nominating president and from the same state as ideologically identical, regardless of their judicial track records.[17]

Only recently have studies begun to develop comprehensive measures of ideology based upon actual judicial performance.[18] Yet such efforts have not yet extended beyond the Supreme Court.[19] Consequently, although there has been a veritable explosion of empirical research about federal judges,[20] with a particular focus on judicial ideology,[21] the most commonly used research measures have been of questionable efficacy in gauging individual judges.[22]

---

[16] *See id.*

[17] *See id.*

[18] *See* Fischman & Law, *supra* note 7, at 162–66.

[19] *See* Cross, *supra* note 9, at 16–17.

[20] A particularly helpful resource for such research is Volume 58, Issue 7 of the *Duke Law Journal*, a special symposium issue on "Measuring Judges and Justice." *See, e.g.*, Jeffrey M. Chemerinsky & Jonathan L. Williams, *Measuring Judges and Justice*, 58 Duke L.J. 1173, 1174 (2009) ("Empirical scholars have begun to train these same tools on the judiciary. They have studied topics ranging from the economic effects of judicial systems to the influence of ideology on judicial decisionmaking."); Frank B. Cross & Stefanie Lindquist, *Judging the Judges*, 58 Duke L.J. 1383, 1384 (2009) ("The rigorous comparative evaluation of federal judges has become a popular topic in academic research . . . ."); Jack Knight, *Are Empiricists Asking the Right Questions About Judicial Decisionmaking?*, 58 Duke L.J. 1531, 1535 (2009) ("Social scientists who study the courts employ an impressive array of statistical and mathematical approaches. This array has grown in variety and sophistication in the last decade."); Gregory C. Sisk, *The Quantitative Moment and the Qualitative Opportunity: Legal Studies of Judicial Decision Making*, 93 Cornell L. Rev. 873, 874 (2008) ("In the past decade, the pace of empirical legal study has quickened, and the publication of empirical studies in law journals has increased. Within just a few short years, empirical study of the law in general, and in particular of the courts, has risen to a level of prominence in American law schools.").

[21] *See, e.g.*, Tonja Jacobi & Matthew Sag, *Taking the Measure of Ideology: Empirically Measuring Supreme Court Cases*, 98 Geo. L.J. 1, 2 (2009) ("A central part of empirical judicial inquiries has been the legal realist and attitudinalist contribution that shows that in most areas of the law, judicial ideology is a significant factor in determining case outcomes."); Gregory C. Sisk & Michael Heise, *Judges and Ideology: Public and Academic Debates About Statistical Measures*, 99 Nw. U. L. Rev. 743, 744 (2005) ("[T]he seclusion of the ivory tower has been breached, as public attention has become increasingly focused upon studies that suggest the influence of ideological and partisan variables on the outcome of court cases.").

[22] *See* Cross, *supra* note 9, at 19; Fischman & Law, *supra* note 7, at 201–02.

This Article attempts to cure this deficiency in empirical research about the federal courts by introducing a new technique for measuring the ideology of judges based upon judicial behavior in the U.S. courts of appeals. Although the actions of the Supreme Court attract greater attention,[23] studying the courts of appeals is significantly more revealing concerning judicial decision making and the changes in federal law.[24] Professors Frank Cross and Stefanie Lindquist recently called for increased research of the courts of appeals by noting that "the circuit court judiciary is probably the single most important level of the federal judiciary in light of its extensive caseload and policy making authority."[25] Further, in an era where it is almost mandatory to first serve on a Court of Appeals before becoming a Supreme Court justice, as was the case with every justice currently on the Court except Justice Elena Kagan, focusing on the appellate level of the federal judiciary can yield important objective information for the nomination process.[26]

The three major difficulties in measuring the ideology of federal judges below the Supreme Court have been: (1) the enormous number of opinions issued by such courts; (2) the inability to accurately code the ideological direction of case outcomes in numerous areas of law; and (3) the high percentage of unanimous decisions at the federal appellate level (which ordinarily defeats efforts to construct a valid baseline for determining ideological differences).[27] The measure advocated herein addresses the first two problems by using "agnostic" measuring

[23] *See* Ashlyn K. Kuersten & Donald R. Songer, Decisions on the U.S. Courts of Appeals 1 (2001) ("The Courts of Appeals are virtually invisible to most Americans . . . . They receive little media coverage because their decisions are often less dramatic than the pronouncements of the Supreme Court . . . .").

[24] *See* Harry T. Edwards & Michael A. Livermore, *Pitfalls of Empirical Studies That Attempt to Understand the Factors Affecting Appellate Decisionmaking*, 58 Duke L.J. 1895, 1904 (2009) (explaining that among the federal courts, the courts of appeals are better to study because "[t]he courts of appeals also hear far more cases each year than does the Supreme Court, have only very limited control over their dockets, and normally sit in panels of three (not en banc)").

[25] Cross & Lindquist, *supra* note 20, at 1385; *see also* David Klein, Making Law in the United States Courts of Appeals 4 (2002) ("The truth, well known but often overlooked in the media and even in serious scholarship, is that lower court judges play a major role in the development of legal doctrine."); Cross, *supra* note 9, at 1–2 ("[T]he circuit courts are much more important [than the U.S. Supreme Court] in setting and enforcing the law of the United States.").

[26] *See* Timothy P. O'Neill, *"The Stepford Justices": The Need for Experiential Diversity on the Roberts Court*, 60 Okla. L. Rev. 701, 702 (2007) ("For the first time in history every justice had been a judge on the U.S. Court of Appeals at the time of appointment to the Supreme Court.").

[27] *See* Edwards & Livermore, *supra* note 24, at 1904; Fischman & Law, *supra* note 7, at 165.

techniques that do not code the ideological direction of individual case outcomes, but instead look at which judges most often vote together.[28] By identifying voting blocs, assessments can be made about the ideologies of the judges that form those blocs.[29] Normally, the third problem—that all but a small percentage of federal appellate judgments are made without a dissenting opinion—makes such agnostic measures useless in such situations.[30] This is because existing agnostic measures cannot infer ideology when a panel of judges is in agreement.[31] This problem is largely eliminated in this study, which treats district judges as hidden fourth members of federal appellate panels. This allows for the incorporation of affirmances and reversals into the ideological measure even when the appellate panel issued a unanimous opinion.

This study measures ideology by determining the degree to which federal appellate judges agree and disagree with their liberal and conservative colleagues at both the appellate and district court levels. Those agreements are indicative in the aggregate of a judge's ideology on the bench. Further, as part of this unique multi-court-level measure, important inter-court factors such as standard of review are incorporated. Not only does this measure integrate review of lower courts, an essential function of the federal judiciary, it also allows for very large sample sizes for individual judges. The metric proposed and applied herein (the "Ideology Score") is based upon opinions derived from or leading to over 30,000 judicial votes issued in 2008 by eleven courts of appeals contained in a unique data set utilized in this study. This Article focuses on 177 appellate-level judges who had an adequate number of interactions with other judges to yield valid inferences. The Article then tests the Ideology Scores against the existing proxy measures to determine which method is superior at predicting judicial votes. Further, this Article examines the connections between judicial ideology and political, social, and demographic characteristics of federal judges by utilizing the newly-proposed measure in a regression analysis.

Based upon the measure outlined in this Article, several important findings emerge. First, the Ideology Scores in this Article offer substantial improvements in predicting civil rights case outcomes over leading

---

[28] *See* Fischman & Law, *supra* note 7, at 156–57. "Agnostic" is a term used by Professors Joshua Fischman and David Law to describe measures that use voting blocs as a means to infer judicial ideology. *See id.* As the term effectively captures the concept, this Article adopts their terminology.

[29] *See id.*

[30] *See id.*

[31] *See id.*

measures of ideology.[32] Second, there were very different levels and heterogeneity of ideology among the judges on the studied circuits.[33] Third, the data did not support the conventional wisdom that Presidents Ronald Reagan and George W. Bush appointed uniquely ideological judges.[34] Fourth, in general judges appointed by Republican presidents were more ideological than those appointed by Democratic presidents.[35] Fifth, attendance at a higher-ranked law school was strongly correlated with liberalism on the bench for appointments of both Republican and Democratic presidents.[36] Sixth, prior work experience in the government (outside of the judiciary) indicated liberal judicial voting.

 This Article is divided into four parts. Part I discusses the meaning of "judicial ideology" and the current methods for measuring the concept.[37] Part II introduces the measure used herein, outlines the study design and methodology, and discusses the advantages associated with the new measure.[38] Part III applies the measure to a newly created data set of 2008 opinions by eleven U.S. courts of appeals and analyzes the results based upon that application.[39] Part IV considers the statistical reliability, validity, and limitations of the study.[40] The Article finishes with some concluding thoughts about judicial ideology and new directions for research. Consistent with the mission of making empirical legal studies more accessible and understandable to a larger audience, this Article avoids empirical research jargon whenever possible and utilizes graphical representation[41] of key measures throughout the Article.[42] Technical details traditionally found in empirical legal studies are

---

[32] *See infra* notes 288–314 and accompanying text.

[33] *See infra* notes 269–279 and accompanying text.

[34] *See infra* notes 266–268 and accompanying text.

[35] *See infra* notes 266–268 and accompanying text.

[36] *See infra* notes 280–281 and accompanying text.

[37] *See infra* notes 43–166 and accompanying text.

[38] *See infra* notes 167–248 and accompanying text.

[39] *See infra* notes 249–281 and accompanying text.

[40] *See infra* notes 282–359 and accompanying text.

[41] *See* Lee Epstein et al., *On the Effective Communication of the Results of Empirical Studies, Part II*, 60 VAND. L. REV. 801, 804–05 (2007) ("[R]esearchers should almost always graph their data and results. . . . Unless the author has a very compelling reason to provide precise numbers to readers, a well-designed graph is a superior choice to a table.").

[42] *See* Lee Epstein et al., *On the Effective Communication of the Results of Empirical Studies, Part I*, 59 VAND. L. REV. 1811, 1814 (2006) ("Most crucially, it seems nearly incontrovertible that moving towards more appropriate and accessible presentations of data will heighten the impact of empirical legal scholarship on its intended audience—be that audience other academics, students, policy makers, lawyers, or judges—not to mention raise the level of intellectual discourse among scholars themselves."); Fischman & Law, *supra* note 7, at 135–36 ("The fact that much of the audience is not methodologically sophisticated

largely located in the footnotes throughout and appendixes at the end of this Article.

I. What Is "Judicial Ideology" and How Can It Be Measured?

Because it is contained in the recesses of a human mind, the ideology of a judge is not directly observable.[43] Ideology is a "latent trait" that can only be empirically studied through surrogate measures.[44] Which proxies a researcher uses can yield very different results and inferences.[45] For example, if a researcher sought to measure a judge's ideology based upon how a judge voted in presidential elections (if such information were available), that might create a very different result than if that researcher looked at how a judge voted in panel decisions concerning intellectual property issues.[46] The choice of proxies is largely based upon implicit and explicit assumptions by a scholar about what "judicial ideology" is.[47] As a result, there are important definitional considerations that anyone researching judicial ideology should address.[48]

A. *Defining and Measuring Judicial Ideology*

Although "ideology" might seem like a concept that everyone understands, it can be difficult to pin down with precision.[49] Notably, in theoretical and empirical discussions of judicial ideology, scholars have rarely defined exactly what they meant when they studied the concept.[50] At its core, "judicial ideology" is a set of political beliefs held by a judge.[51] Nevertheless, that broad definition does not identify which beliefs are truly ideological or how they should be measured.[52]

---

makes it all the more crucial that we do so. As a research community, we must cultivate and convey a better understanding of methods for measuring judicial ideology if we are to succeed in convincing others of the validity of our work.").

[43] *See* Fischman & Law, *supra* note 7, at 143–44.

[44] *See id.*

[45] *See id.* at 166–67.

[46] *See* Matthew Sag et al., *Ideology and Exceptionalism in Intellectual Property: An Empirical Study*, 97 Cal. L. Rev. 801, 847 (2009). Although intellectual property might seem rife with apolitical questions, a recent study demonstrated that ideology predicts a great number of judicial decisions in such cases. *See id.*

[47] *See* Fischman & Law, *supra* note 7, at 137–39.

[48] *See id.*

[49] *See id.* at 135.

[50] *See id.* ("Empirical studies routinely purport to measure ideology without specifying what is meant by 'ideology . . . .'").

[51] *See id.* at 137 ("An ideology is, in a literal sense, a collection or system of ideas.").

[52] *See id.* at 137–38.

Thus far, the primary scholarly impetus for measuring judicial ideology has been to determine the degree to which ideology affects decision making.[53] That task illustrates why the broad definition has not generally sufficed.[54] There has been considerable debate among scholars and judges concerning the subject of ideological voting by judges.[55] On one side have been the strong "attitudinalists" who have contended that ideology is the core determining factor in judicial decision making.[56] In contrast, some scholars and judges, such as Judge Harry T. Edwards,[57] have argued that formalist interpretation of law, and not ideology, resolves the large majority of legal disputes.[58] Others, such as Judge Richard Posner, have argued that strategic, institutional goals (such as career or reputational advancement) offer the best explanation for judicial behavior.[59] The broad definition of ideology does not yield measures to resolve these disputes because it lacks any content describing legal, ideological, and strategic decisions. Consequently, each side of the judicial decision-making debate can argue that the re-

---

[53] *See* Fischman & Law, *supra* note 7, at 138.

[54] *See id.*

[55] *See id.* at 135 ("Without seeking to deny that ideology plays a significant role in judicial decision-making, well-informed observers have nevertheless raised reasonable criticisms about the manner in which empirical scholars have tackled the subject . . . . Empirical studies routinely purport to measure ideology without specifying what is meant by 'ideology,' or taking care to measure 'ideology' in a way that will not invite a host of objections.").

[56] *See* Jeffrey A. Segal & Harold J. Spaeth, The Supreme Court and the Attitudinal Model 56 (1993) ("This [attitudinal] model holds that the Supreme Court decides disputes in light of the facts of the case vis-à-vis the ideological attitudes and values of the justices.").

[57] *See* Harry T. Edwards, *Collegiality and Decision Making on the D.C. Circuit*, 84 Va. L. Rev. 1335, 1336 (1998); Edwards & Livermore, *supra* note 24, at 1904–05; Harry T. Edwards, *Public Misperceptions Concerning the "Politics" of Judging: Dispelling Some Myths About the D.C. Circuit*, 56 U. Colo. L. Rev. 619, 625 (1985); Harry T. Edwards, *The Effects of Collegiality on Judicial Decision Making*, 151 U. Pa. L. Rev. 1639, 1656 (2003) [hereinafter Edwards, *The Effects of Collegiality*].

[58] *See* Fischman & Law, *supra* note 7, at 133 ("There remains deep skepticism in legal circles toward interdisciplinary empirical scholarship aimed at capturing the impact of ideology on judicial behavior.").

[59] *See* Richard A. Posner, How Judges Think 29 (2008) ("The strategic theory of judicial behavior (also called the positive political theory of law) . . . hypothesizes that judges do not always vote as they would if they did not have to worry about the reactions to their votes of other judges (whether their colleagues or the judges of a higher or a lower court), legislators, and the public.").

sults derived from the broad definition fit with its preferred hypothesis.[60]

Because of the general inadequacy of the simplistic meaning of ideology, scholars have sometimes used "judicial ideology" in a similar manner, but at other times, have employed the term in significantly different ways.[61] A common conception of ideology is that it is synonymous with politics and antithetical to law.[62] What that means in empirical research, however, is ambiguous. If a judge is liberal, does that mean he or she favors particular parties in legal cases, such as criminal defendants or ethnic minorities?[63] Or does it mean that the judge utilizes an interpretative methodology that is not concerned with the historical understanding of the legal provisions being reviewed?[64] Or is a liberal judge focused on achieving certain policy ends such as equality, greater regulation of corporations, and access to government services?[65]

Although it might seem that all of those possibilities characterize the ideology of a liberal judge, each indicates a preference for very different ways of measuring the concept. For example, if the emphasis of research is on favored or disfavored types of parties (e.g., criminal defendants), the researcher would be concerned with the prevailing parties before the judges.[66] If the interpretative method is the touchstone of liberalism, however, the reasoning of the opinion will be the focus of coding.[67] If the scholar believes that a set of policy goals defines what liberalism is for a judge, some combination of the prevailing party and the reasoning of the opinion will have to be analyzed within a larger

---

[60] *See* Fischman & Law, *supra* note 7, at 147 ("The ongoing debate over the relative merits of the legal, attitudinal, and strategic models persists in part because the behavior that we observe can often be explained in more than one way.").

[61] *See id.* at 137 ("Scholars have used the term 'ideology' in a bewildering variety of ways, often without even attempting to define it. The result is that ideology is in the eye of the beholder: what one observer might call ideological behavior, another might call principled judging, and vice versa.").

[62] *See id.* at 138.

[63] *See id.*

[64] *See id.*

[65] *See id.*

[66] *See* Fischman & Law, *supra* note 7, at 141–42 ("Alternatively, the term 'ideological' could describe a tendency to favor or disfavor certain types of parties—criminal defendants, police officers, corporations, members of ethnic or religious minorities, the disabled, and so forth.").

[67] *See id.* ("To say that a certain type of judicial behavior is 'ideological' need not mean that it is ideological in a *political* sense: one might, for example, characterize adjudication that relies heavily upon logical deduction from formal rules as narrowly 'legal,' whereas adjudication driven by ideas about the role of law and the responsibilities of judges might by contrast be characterized as both 'legal' and 'ideological' in character.").

social context.[68] Further, although the focus on parties, for example, in criminal cases would yield a robust data set, a researcher concerned with policy goals might discount an inordinate number of the cases as being irrelevant to the articulated liberal goals.[69] The nature of the measure and the way data is collected are dictated by the definition adopted or implied by the empirical researcher.[70]

Rather than engaging in the highly subjective and problematic determination of the meaning of ideology and the proper measurement to determine its effects, this Article takes a different approach. Instead of focusing on identifying which decisions were ideological and then determining the aggregate degree to which ideology drove a judge's vote, this Article accepts as a basic premise that at least some body of decisions were guided by a judge's ideology. Even the most ardent supporters of strategic and legal models of decision making acknowledge that a portion of decisions are best explained by ideology.[71] Further, this premise is consistent with prior research of the federal judiciary.[72] The focus of the measure herein is gauging the judges in the aggregate based upon situations when their ideologies surface in their voting behavior. This study compares judges to determine which ones are more conservative or liberal relative to their colleagues based upon whom they most often vote with and against.

The result is a set of scores for judges ranging from the most liberal to the most conservative judges serving on the U.S. courts of appeals, but that scale is not tied to some larger objective measure—it is

---

[68] *See id.*

[69] *See id.* ("It might be considered 'ideological,' for example, for judges to seek to advance a particular policy outcome . . . .").

[70] *See id.* ("It is not possible for this Article to settle the correct meaning of the term 'ideology.' How researchers should or will use the term will inevitably depend upon the purposes that they have in mind.").

[71] *See, e.g.*, Edwards & Livermore, *supra* note 24, at 1899 ("Because it is undisputed that some cases admit of discretion in the exercise of appellate decisionmaking, scholars and commentators sometimes contend that judges must be influenced in their decisionmaking by their personal political or ideological predilections. This may happen at times."); Sisk & Heise, *supra* note 21, at 752–53 ("The real debate . . . was not whether ideology or extra-legal factors matter in judging, but how much, how often, in what instances, and whether they substantially undermine the proper functioning of the judiciary.").

[72] *See, e.g.*, Frank B. Cross, *Decisionmaking in the U.S. Circuit Courts of Appeals*, 91 Cal. L. Rev. 1457, 1479–82 (2003) (reviewing the empirical evidence in support of ideological decision making in federal courts below the U.S. Supreme Court); Sisk & Heise, *supra* note 21, at 746 ("The growing body of empirical research on the lower federal courts . . . reveals that ideology explains only a relatively modest part of judicial behavior and emerges on the margins in controversial and ideologically contested cases.").

entirely relative.[73] Because of the nature of the measure used in this study, the definitional problems are less significant than in other empirical work.[74] This Article uses the broad definition described above because there is no need to outline the actual content of any judge's ideology. As will be further explained in Part II, by making prior assessments of particular judges being liberal and conservative, there is no need to describe the particulars of their belief systems.[75] It is merely enough to measure the degree to which the studied judges agree or disagree with their liberal and conservative colleagues.

B. *Measures of the Judicial Ideology of Federal Judges*

There have been many attempts by legal academics and political scientists to gauge the ideology of individual judges at different levels of the judiciary.[76] Because of the available resources, most of these studies have focused on justices of the United States Supreme Court.[77] Certain techniques have worked very well in studying the highest court; however, they are not as useful in analyzing other courts.[78] Although the measures of ideology can be categorized in a variety of manners, perhaps the easiest way to appreciate the real differences in techniques is to consider three major approaches: case outcome coding, external proxies, and agnostic measures.

1. Measuring Judicial Ideology Through Case Outcome Coding

Using case outcomes to determine judicial ideology requires scholars to assess a judge's votes or opinions based upon the ideological direction of the decisions the judge has made.[79] Cases are typically coded using a binary construction where a particular case outcome is either entirely "liberal" or "conservative."[80] Based upon a judge's votes or opin-

---

[73] The scale is limited to judges with sufficient sample sizes to create valid statistical inferences. *See infra* notes 190–191 and accompanying text.

[74] *See* Fischman & Law, *supra* note 7, at 138.

[75] *See infra* notes 243–248 and accompanying text.

[76] *See, e.g.*, Klein, *supra* note 25, at 4; Cross & Lindquist, *supra* note 20, at 1385.

[77] *See, e.g.*, Segal & Spaeth, *supra* note 56, at xv; Jacobi & Sag, *supra* note 21, at 4.

[78] *See* Fischman & Law, *supra* note 7, at 152.

[79] *See id.* at 156–62.

[80] *See id.* at 156 ("Most empirical studies on the subject of judicial ideology rely on some sort of dichotomous coding scheme, in which observable judicial actions—typically a vote or decision of some kind—are coded as 'zero' or 'one,' depending on whether they are, in some rough sense, 'liberal' or 'conservative.'"); Jacobi & Sag, *supra* note 21, at 6–7 ("The primary form of categorization is fairly basic . . . . The standard means for categoriz-

ions in a sufficient sample of cases, a judge's ideology can be identified.[81] There is intuitive appeal to this methodology because it is how a lay audience, believing that the labels "conservative" and "liberal" are easily applied to particular decisions, would presumably assess a judge.[82]

Because of the enormous work effort required for such a project, scholars using case outcome methodologies have primarily relied upon two pre-existing data sets.[83] For studies of the Supreme Court, scholars have typically utilized Harold Spaeth's United States Supreme Court Judicial Database (the "Spaeth Database").[84] For the U.S. courts of appeals, the Appeals Court Database (the "Songer Database"), which includes more than 18,000 opinions from 1925 to 1996,[85] has become the data set of choice.[86] There is no similar database for federal district courts, and research at that level has been extremely limited.[87] The Spaeth Database is comprehensive in regards to opinions issued by the Court, so it should not be surprising that most of the empirical research of judicial ideology has been focused on the Supreme Court.[88] Because the Songer Database covers such a lengthy time period with limited sampling, however, it is virtually impossible to make valid ideology assessments of individual judges from this source. Even studies that have focused on discrete courts and issues, such as the examination of administrative deference in the D.C. Circuit by Professors Emerson Tiller and Frank Cross, or Professor Richard Revesz's examination of environmental decisions in the D.C. Circuit, have not provided individual judge ideology measures.[89] Overall, in the federal courts below the

---

ing case outcomes is to use Spaeth's United States Supreme Court Judicial Database . . . coding, which categorizes cases as 'liberal' or 'conservative.'").

[81] *See* Jacobi & Sag, *supra* note 21, at 7.

[82] *See* Fischman & Law, *supra* note 7, at 163.

[83] *See* KUERSTEN & SONGER, *supra* note 23, at 241; Fischman & Law, *supra* note 7, at 161.

[84] *See* Jacobi & Sag, *supra* note 21, at 8; Harold J. Spaeth, THE SUPREME COURT DATABASE (last updated Aug. 26, 2010), http://supremecourtdatabase.org.

[85] KUERSTEN & SONGER, *supra* note 23, at 241; Donald R. Songer, APPEALS COURT DATABASE, http://people.cas.sc.edu/songer/ (follow hyperlinks next to "Appeals Court Database") (last visited Aug. 10, 2010).

[86] *See* Cross, *supra* note 72, at 1498.

[87] *See* Edwards & Livermore, *supra* note 24, at 1899–1900.

[88] *See* KUERSTEN & SONGER, *supra* note 23, at 241.

[89] *See* Frank B. Cross & Emerson H. Tiller, *Judicial Partisanship and Obedience to Legal Doctrine: Whistleblowing on the Federal Courts of Appeals*, 107 YALE L.J. 2155, 2173 (1998); Richard L. Revesz, *Environmental Regulations, Ideology, and the D.C. Circuit*, 83 VA. L. REV. 1717, 1763–64 (1997). David Law's asylum study, however, allowed for individual assessments in the Ninth Circuit. David S. Law, *Strategic Judicial Lawmaking: Ideology, Publication, and Asylum Law in the Ninth Circuit*, 73 U. CINN. L. REV. 817, 860 (2005).

Supreme Court, there has been no comprehensive measure of individual judge ideology using case outcome coding.[90]

Case outcome coding is susceptible to two unique problems.[91] First, it is extremely labor-intensive because every case, even ones with contradictory or indeterminate indicators, must be coded by ideology.[92] This burden is particularly high if a scholar wants to make individual assessments of a large number of judges, because a sufficient sample will be required for each judge.[93] When data set creation is too difficult for a researcher, he or she is usually forced to shrink the breadth of the data set.[94] Limiting the scope of the data set can diminish the inferences that can be drawn from a study as well as increase standard errors.[95]

Second, case outcome coding relies on subjective coding that is often imprecise.[96] As Professors Tonja Jacobi and Matthew Sag recently observed, "[T]he last four decades of empirical scholarship have proceeded without a sophisticated objective measure of case outcomes."[97] Because of its high profile, the Spaeth Database has been targeted by critics because of the problems of subjective and imprecise coding.[98] Professor Anna Harvey compared the ideology coding decisions made in the Spaeth Database with the roll call voting in Congress when federal legislation was reviewed.[99] Her results indicated little correlation between the ideology of the congressional votes and the Spaeth Database coding.[100] Judge Richard Posner and Professor William Landes examined both the Spaeth Database and Songer Database for ideological coding and found numerous discrepancies between their coding and that contained in the databases.[101] Further, because case coding relies upon a wholly binary construction of the concept, it lacks signifi-

---

[90] *See* Fischman & Law, *supra* note 7, at 156–62.

[91] *See id.*

[92] *See id.*

[93] *See id.*

[94] *See id.* at 157–58.

[95] *See* ROBERT M. LAWLESS ET AL., EMPIRICAL METHODS IN LAW 425 (2010) (defining a standard error as "an estimate of the error between a sample statistic and the true value in the population"); Fischman & Law, *supra* note 7, at 157–58.

[96] *See* Jacobi & Sag, *supra* note 21, at 7.

[97] *Id.*

[98] *See* Carolyn Shapiro, *Coding Complexity: Bringing Law to the Empirical Analysis of the Supreme Court*, 60 HASTINGS L.J. 477, 488–500 (2009).

[99] *See* Anna Harvey, What Makes a Judgment "Liberal"? Coding Bias in the United States Supreme Court Judicial Database 17 (June 15, 2008) (unnumbered working paper), *available at* http://ssrn.com/abstract=1120970.

[100] *See id.* at 19.

[101] *See* William M. Landes & Richard A. Posner, *Rational Judicial Behavior: A Statistical Study*, 1 J. LEGAL ANALYSIS 775, 784 (2009).

cant nuance in particular cases.[102] For example, because of its simplistic design, a vote for striking down a restriction on access to abortion is coded as having the same ideology and importance as a plaintiff victory in a securities derivative suit.[103]

2. Measuring Judicial Ideology Through External Proxies

External proxy metrics determine judicial ideology based upon responses to judges' work by others, without analyzing the content of the opinions of those judges.[104] The external audience can be composed of other judges, legislators, agencies, the media, or members of the general public.[105] The reception of a judge's decisions can give insight into the judge's incentives as well as reasoning.[106] Nonetheless, external proxy metrics, by their very nature, do not directly measure the qualities contained within the judicial opinions issued.[107]

Outside of ideology studies, the most notorious recent example of the use of external proxy measures was the "tournament of judges" model created by Professors Stephen Choi and G. Mitu Gulati.[108] That study evaluated judges on the U.S. courts of appeals based upon, among other measures, two external proxy metrics: citations to opinions of particular judges, and the frequency with which a judge reached a different judgment than members of his or her party on the same panel.[109] Similarly, Professors Jeffrey Berger and Tracey George completed a study of judicial entrepreneurs at the federal appellate level

---

[102] *See* Jacobi & Sag, *supra* note 21, at 6 ("Due to this unreliability, scholars more commonly rely on objective, but highly simplified, means of categorizing or labeling cases. These categories are typically dichotomous—for example, a case is either pro-plaintiff or pro-defendant—and thus do not provide as meaningful a form of measurement as a continuous variable would. A continuous measure—one that permits infinite variation along a scale—would allow cases to be judged accurately by comparing the score with a substantive, comparative assessment of similar cases.").

[103] *See id.*

[104] *See* Fischman & Law, *supra* note 7, at 166–76.

[105] *See* POSNER, *supra* note 59, at 29.

[106] *See id.*

[107] *See* Fischman & Law, *supra* note 7, at 166–67.

[108] *See* Stephen Choi & Mitu Gulati, Essay, *A Tournament of Judges?*, 92 CAL. L. REV. 299, 305–13 (2004) [hereinafter Choi & Gulati, *Tournament*]; Stephen J. Choi & G. Mitu Gulati, *Choosing the Next Supreme Court Justice: An Empirical Ranking of Judge Performance*, 78 S. CAL. L. REV. 23, 33–34 (2004); Stephen J. Choi & G. Mitu Gulati, *Mr. Justice Posner? Unpacking the Statistics*, 61 N.Y.U. ANN. SURV. AM. L. 19, 29 (2005).

[109] Choi & Gulati, *Tournament, supra* note 108, at 305–13.

that focused on citations of judges by the Supreme Court.[110] Professors David Klein and Darby Morrisroe measured prestige of federal appellate judges through the external metric of citations by other judges.[111] In a subsequent study, Klein incorporated the metric of whether a judge established a new rule that was followed by later courts.[112] Professors Frank Cross and Stefanie Lindquist have recently published a study that focused on reversal rates upon appeal to the Supreme Court.[113]

In the context of judicial ideology, external proxies have been the most popular method of assessing the political beliefs of particular judges.[114] Indeed, the most common measure used for the ideology of federal judges by scholars has been the external proxy of the political party of the appointing president.[115] Although research has shown that the party of the appointing president has a statistically significant effect on the decisions of judges on the U.S. courts of appeals in some areas of law,[116] studies have not found a similar relationship at other court levels.[117]

Segal-Cover scores and NOMINATE scores have become commonly used metrics for judicial ideology in certain circumstances.[118] Segal-Cover scores apply strictly to Supreme Court justices and are based upon newspaper editorials from four major newspapers at the time of nomination.[119] Although Segal-Cover scores have been subject to praise[120] and

---

[110] *See* Jeffrey A. Berger & Tracey E. George, *Judicial Entrepreneurs on the U.S. Courts of Appeals: A Citation Analysis of Judicial Influence* 1, 3 (Vanderbilt U. L. Sch. L. & Econ., Working Paper No. 05-24, 2005), *available at* http://ssrn.com/abstract=789544.

[111] *See* David E. Klein & Darby Morrisroe, *The Prestige and Influence of Individual Judges on the U.S. Courts of Appeals*, 28 J. Legal Stud. 371, 379 (1999).

[112] *See* Klein, *supra* note 25, at 51.

[113] *See* Cross & Lindquist, *supra* note 20, at 1405 ("We do not suggest that the Supreme Court review measure is the sole test for circuit court judicial quality or even a 'gold standard' measure. We contend only that it is an important metric that has not been previously considered and that may yield insight into the quality of circuit court judges.").

[114] *See* Fischman & Law, *supra* note 7, at 166–67.

[115] *See id.* at 167–68.

[116] *See* Cass R. Sunstein et al., *Ideological Voting on Federal Courts of Appeals: A Preliminary Investigation*, 90 Va. L. Rev. 301, 303 (2004).

[117] *See* Orley Ashenfelter et al., *Politics and the Judiciary: The Influence of Judicial Background on Case Outcomes*, 24 J. Legal Stud. 257, 276 (1995).

[118] *See* Keith T. Poole & Howard Rosenthal, Congress: A Political-Economic History of Roll Call Voting 27–29 (1997) (discussing NOMINATE Scores, which use roll-call data from Congress to estimate the ideological ideal points of legislators); Jeffrey A. Segal & Albert D. Cover, *Ideological Values and the Votes of U.S. Supreme Court Justices*, 83 Am. Pol. Sci. Rev. 557, 558 (1989) (calculating Segal-Cover Scores).

[119] Segal & Cover, *supra* note 118, at 559.

[120] *See* Fischman & Law, *supra* note 7, at 184–85.

criticism,[121] a similar approach at the lower federal courts is simply impossible because of the limited, and event-driven coverage of other federal judges.[122]

NOMINATE Scores have been subsequently updated and modified into Common Space Scores. Common Space Scores have become increasingly popular as an alternative to the appointing president proxy measure.[123] Common Space Scores integrate the politics of the appointing president with the home-state senator(s) of the judicial nominee.[124] The metric relies on the notion that senatorial courtesy is an important factor in the selection of federal judges.[125] Common Space Scores move beyond the simple binary construction of the appointing president while accounting for a larger political context at the time of a judge's nomination.[126] Nonetheless, recent research by Joshua Fischman and David Law analyzed the efficacy of Common Space Scores in predicting outcomes in Ninth Circuit asylum cases and found only marginal improvements over the appointing president measure.[127]

As is made clear by the above examples, external proxy scores are quite varied; however, they share several shortcomings.[128] First, similar to case outcome coding, external proxies often rely on either subjective coding decisions or imprecise coding.[129] In the case of the party of the appointing president, ideology is reduced to a binary concept because of America's two-party political system.[130] Further, every judge appointed by a president is treated identically.[131] Even presidents with very different ideologies would have their nominated judges labeled as ideological matches using the appointing president proxy if they were of the same political party.[132] Thus, the appointing president measure fails to acknowledge any gradations along the political spectrum.[133] Al-

---

[121] *See id.*

[122] *See* KUERSTEN & SONGER, *supra* note 23, at 1.

[123] Fischman & Law, *supra* note 7, at 173.

[124] Lee Epstein & Gary King, *The Rules of Inference*, 69 U. CHI. L. REV. 1, 83–84, 95–96 (2002).

[125] *See* Giles, *supra* note 11, at 634–35.

[126] *See id.*

[127] *See* Fischman & Law, *supra* note 7, at 167–68.

[128] *See* CROSS, *supra* note 9, at 20.

[129] *See id;* Fischman & Law, *supra* note 7, at 170.

[130] *See* Fischman & Law, *supra* note 7, at 170.

[131] *See id.*

[132] *See id.*

[133] *See id.* at 171 ("Party of appointment can be used to identify aggregate differences between appointees of the two parties, but it cannot distinguish between moderate and extreme judges.").

though adding the politics of the senators into the judges' ideologies creates greater variation in scoring, it is unclear if such distinctions actually reflect the ideologies of those judges more accurately than the party of the appointing president alone.[134] Further, although the Common Space Scores create a greater range of scores, the appointing president's political party still entirely controls whether a judge is on the liberal or conservative half of the spectrum.[135]

Second, because an external metric relies on responses by other parties, the metric is only as good as those other parties.[136] So, for example, if the U.S. Supreme Court were less competent than the lower appellate courts at analyzing certain legal issues,[137] a measure of judicial quality based upon Supreme Court reversals would actually be inversely related to quality.[138] Similarly, if judges cited other judges' opinions based solely upon reputation, personal relations, or showiness in opinion writing, then the citations would fail to effectively demonstrate quality.[139] Further, feedback loops can develop that exacerbate those effects to a larger extent over time.[140] In the context of ideology studies, Segal-Cover scores are particularly susceptible to this criticism because they rely on editorial writers who may be driven by partisan goals or who lack any real insight into the targets of their writing.[141] Common Space

---

[134] *See id.* at 167–68.

[135] *See id.*

[136] *See* Cross & Lindquist, *supra* note 20, at 1400–03.

[137] *See* David L. Schwartz, *Practice Makes Perfect?: An Empirical Study of Claim Construction Reversal Rates in Patent Cases*, 107 MICH. L. REV. 223, 265–66 (2008). Although there is a general assumption that higher courts are more competent, such a proposition has not been empirically supported. *See id.* In a different area of federal law, Schwartz argued that because federal district courts do not lower their reversal rates in patent claim construction cases, the problem might be that the Federal Circuit has repeatedly made a mess of the relevant controlling law. *See id.* ("Alternatively, the problem may stem from potential inconsistencies in the Federal Circuit's methodology or ideology. If such inconsistencies exist, then they prevent parties from predicting accurate claim constructions prior to appellate review.").

[138] *See* Cross & Lindquist, *supra* note 20, at 1403. Cross and Lindquist contend that "this criticism inevitably undermines every attempt to measure judicial quality and thus is not dispositive." *Id.*

[139] *See* Daniel A. Farber, *Supreme Court Selection and Measures of Past Judicial Performance*, 32 FLA. ST. U. L. REV. 1175, 1178–79 (2005) (discussing factors that influence how often a judge's opinions are cited, including the author's reputation, personal relationships between judges, and some degree of luck); Kate O'Neill, *Rhetoric Counts: What We Should Teach When We Teach Posner*, 39 SETON HALL L. REV. 507, 507–09 (2009) (finding a connection between Judge Posner's opinions and the stylistic nature of case book citations).

[140] *See* Farber, *supra* note 139, at 1177.

[141] *See* Lee Epstein & Carol Mershon, *Measuring Political Preferences*, 40 AM. J. POL. SCI. 261, 281–84 (1996).

Scores similarly rely on a senatorial courtesy norm that might not be used regularly by certain presidents or in specific situations.[142]

Third, the leading proxy measures all rely on factors at the time of nomination instead of the behavior of judges after their appointment.[143] This means that the ratings are forever trapped in time and do not account for changing circumstances.[144] If a judge turns out to have a different ideology than expected (e.g., Justice Souter), the external proxies treat the judge as though the expectations were met.[145] Further, if a judge changes ideology over time (e.g., Justice Blackmun), the measures do not recognize the change.[146]

3. Measuring Judicial Ideology Through Agnostic Measures

Agnostic measures identify voting blocs in a large number of cases to determine which judges are most often aligned.[147] Thus, agnostic measures do not rely on the coding of particular cases on subjective criteria or on factors external to the judiciary.[148] If a judge more often votes with liberals and against conservatives, then an agnostic measure-based study would conclude that the judge is liberal.[149] The rate of agreements and disagreements determines the intensity of that ideology.

Notably, in order for an agnostic measure to work effectively, the researcher must be able to make some prior assessments of particular judges that embody the qualities studied.[150] For example, if a scholar wanted to determine whether persons who wore the same color shirts were more inclined to sit near each other in class, the scholar would need to have prior assessments of what shirts in the class constituted red, blue, or green. In the context of judicial ideology studies, there must be some ability to identify who are some conservative and liberal

---

[142] *See* Giles, *supra* note 11, at 628.

[143] *See* Fischman & Law, *supra* note 7, at 173.

[144] *See id.*

[145] *See* Lee Epstein et al., *Ideological Drift Among Supreme Court Justices: Who, When, and How Important?*, 101 Nw. U. L. Rev. 1483, 1519–26 (2007) (finding an ideological drift among U.S. Supreme Court justices over time).

[146] *See id.*

[147] *See* Fischman & Law, *supra* note 7, at 162–66.

[148] *See id.* at 162.

[149] *See id.*

[150] *See* Andrew D. Martin & Kevin M. Quinn, *Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953–1999*, 10 Pol. Analysis 134, 139–40 (2002).

judges.[151] The judge need not be liberal or conservative all of the time, nor the most liberal or conservative judge.[152] It is simply enough to be confident that a significant portion of the identified judges belong on either half of the political spectrum.[153]

Of the existing ideology measures, the leading one applied to the Supreme Court is agnostic: Martin-Quinn scores.[154] Professors Andrew Martin and Kevin Quinn constructed a model whereby voting alignments of justices in non-unanimous opinions since 1953 served as the basis for ideology scores.[155] They then used a simulation technique known as Markov Chain Monte Carlo simulation to provide a probability distribution of each of the justices' ideological ideal points.[156] As detailed in Part II, this Article utilizes a method akin to the Martin-Quinn scores, but does not use Markov Chain Monte Carlo methods due to software limitations and inconsistencies with the nature of appellate and trial court interactions.[157]

As currently constructed, agnostic measures have two unique disadvantages.[158] First, such measures exhibit data set breadth problems because neither unanimous panel opinions nor opinions made outside of a panel structure (e.g., those at the federal district court level) can be effectively integrated into the measure.[159] When panel judges are in

---

[151] *See id.*

[152] *See id.*

[153] *See id.*

[154] *See id.*

[155] *See id.*

[156] *See* Martin & Quinn, *supra* note 150, at 139–40.

[157] *See id.* There were four reasons that this Article did not use Markov Chain Monte Carlo methods. First, the software utilized by Martin and Quinn, MCMCpack only allows for data to be entered as zeroes, ones, and blanks. *See id.* at 141. Although that works well for the Supreme Court, this model requires opinions to be weighted on a continuous scale. Second, the Markov Chain Monte Carlo method, as currently constructed, does not utilize unanimous opinions. *See id.* at 137 n.3. One of the substantial improvements offered by the model used herein is to add those cases into the model. The first two problems, along with a few other problems, afflict the other leading Markov Chain Monte Carlo methods software, WinBUGS. *See* Law, *supra* note 89, at 857 n.127. Third, unlike the Supreme Court, just three or four (if the district court judge is included) judges review a case at a given time. *See* Edwards, *The Effects of Collegiality, supra* note 57, at 1656. This makes the points of interaction among the judges occur with far less frequency. *See id.* This results in significant difficulties in achieving convergence, which is a necessary part of the Markov Chain Monte Carlo methods. Fourth, one of the primary reasons for using such a model—the ability to deal with small samples using Bayesian statistical techniques—was not a particular concern in this study because of the size of the data set. *See* Martin & Quinn, *supra* note 150, at 135.

[158] *See* Fischman & Law, *supra* note 7, at 165; Martin & Quinn, *supra* note 150, at 139–40.

[159] *See* Fischman & Law, *supra* note 7, at 165; Martin & Quinn, *supra* note 150, at 137 n.3.

agreement, the researcher cannot determine the ideological direction of the decision because the ideology of the case outcome is not coded using such a method.[160] The agreement of the three judges, thus, cannot be labeled as an agreement indicative of liberalism or conservatism.[161] This problem is even more significant if, as Joshua Fischman has found, the courts studied exhibit norms of consensus that mask underlying ideological disagreement.[162] Further, trial judges are not part of voting blocs, so agnostic measures have been useless as to such judges.[163]

Second, agnostic measures cannot exist independent of some other assessment of ideology.[164] Agnostic measures need strong prior assumptions about at least some judges who can be characterized as liberal or conservative. Without such reliable and valid assumptions, an agnostic measure will still yield results, but it will be unclear what has been measured.[165] In the context of ideology, however, because there is fairly wide agreement about some prior assessments concerning judges who are liberal and conservative, agnostic measures can overcome this potential pitfall.[166]

II. Measuring Judicial Ideology by Agreement and Disagreement Among Federal Judges at Multiple Levels

The model used herein relies on the idea that like-minded judges will vote together more often. If there is enough data concerning agreements and disagreements, groups of judges with similar belief systems can be identified. With some prior assessments of which judges are more likely to be conservative or liberal, the belief systems can be mapped onto the traditional left-right political spectrum.[167] As previously noted, this model integrates district court voting into the determination of appellate judge ideology.[168]

This measure not only expands ideology measurement beyond the U.S. Supreme Court, but it also allows for a substantially more robust

---

[160] *See* Fischman & Law, *supra* note 7, at 165.

[161] *See id.*

[162] *See id.*; Joshua B. Fischman, *Estimating Preferences of Appellate Judges: A Model of "Consensus Voting"* 3–4 (Va. Law & Econ., Research Paper No. 2009-14; Va. Pub. Law & Legal Theory, Research Paper No. 2009-21, 2009), *available at* http://ssrn.com/abstract=1361348.

[163] *See* Fischman & Law, *supra* note 7, at 165.

[164] *See* Martin & Quinn, *supra* note 150, at 139–40.

[165] *See id.*

[166] *See* Fischman & Law, *supra* note 7, at 162–63.

[167] *See id.*

[168] *See supra* notes 31–32 and accompanying text; *infra* notes 189–191 and accompanying text.

model that incorporates the full range of decisions made by federal judges. There are many aspects of the modern Supreme Court that make it a highly atypical court in terms of judicial ideology.[169] Primarily, because of the ever-shrinking Supreme Court docket, it is a less-than-ideal institution to study empirically.[170] With approximately seventy-five cases reviewed a year, the population sizes are very small (particularly using existing measures),[171] coverage of different areas of law is minimal,[172] and the self-selected docket does not necessarily provide a random sample of litigation in the United States even while overlooking the selection effects of the litigants involved.[173] In contrast, the U.S. courts of appeals review an enormous number of cases over a wide area of law,[174] are constrained by the Supreme Court, en banc review, and other panel decisions,[175] and judges on those courts can be assessed based upon performance before being elevated to the Supreme

---

[169] *See* Edwards & Livermore, *supra* note 24, at 1904 ("There are many empirical studies devoted to the decisions of the Supreme Court. However, because of the Court's unique status and operating procedures, it is difficult to draw broad conclusions about decision-making in the federal courts of appeals from studies of the Supreme Court.").

[170] *See* Cross, *supra* note 9, at 2 ("The Supreme Court now decides only seventy-five cases a year . . . .").

[171] *See id.*

[172] *See id.*; Richard A. Posner, *Judicial Behavior and Performance: An Economic Approach*, 32 Fla. St. U. L. Rev. 1259, 1273 (2005) ("[T]he Supreme Court reviews only a minute percentage . . . of court of appeals decisions. Entire fields of law are left mainly to the courts of appeals to shape.").

[173] *See* Jonathan Masur, *Judicial Deference and the Credibility of Agency Commitments*, 60 Vand. L. Rev. 1021, 1058 n.162 (2007) ("The drawing of conclusions based on this collection of Supreme Court cases is obviously fraught with peril because the sample is heavily influenced by selection effects. Only certain types of cases reach the Supreme Court (or are even litigated in the first place), and unquestionably ambiguous and unambiguous statutes are likely filtered out at lower levels of the process."); Richard L. Revesz, *Empirical Research and the Goals of Legal Scholarship: A Defense of Empirical Legal Scholarship*, 69 U. Chi. L. Rev. 169, 173–74 & n.25 (2002) ("[A]ny empirical technique that does not control for case-selection effects is biased against finding evidence consistent with the hypothesis that judges vote ideologically."); Frederick Schauer, *Constitutional Positivism*, 25 Conn. L. Rev. 797, 824 (1993) ("[G]iven the existing indeterminacy and moral reference of so much of the constitutional text, and given the way in which the selection effect concentrates in the Supreme Court cases for which the narrowly legal materials do not generate an answer, the actual process of constitutional decisionmaking in the Supreme Court of the United States is one in which Dworkin is much more right than he is wrong.").

[174] *See* Cross, *supra* note 25, at 2 ("[T]he circuit courts resolve more than fifty thousand cases a year.").

[175] *See* Charlie Savage, *Uncertain Evidence for "Activist" Label on Sotomayor*, N.Y. Times, June 20, 2009, at A10 ("Supreme Court justices have a freer hand than appeals court judges.").

Court.[176] This last aspect of U.S. courts of appeals judges is particularly notable given the strong trend to select Supreme Court justices from the federal appellate ranks.[177]

The value of the agnostic measure used in this study is especially clear when comparing it to the alternatives. Whereas ideological case outcome coding might not be able to identify, for example, whether a particular tort action or contract dispute was resolved in a "liberal" manner, agnostic coding should identify trends in the aggregate, insofar as conservatives and liberals reach different conclusions in such cases.[178] If there are no ideological dividing lines in those cases, then the agnostic measure would similarly ignore those decisions because they were not indicative of judicial ideology.[179] Importantly, by adding district court votes to the model of appellate court judge ideology, the measure is able to harvest data from unanimous appellate panel opinions.[180] And, as explained below, information about an appellate judge's ideology is gained even when a unanimous appellate panel affirms a district court judgment.[181] Also, the measure greatly expands the number of comparative references for appellate judges from their appellate judge colleagues to all those trial judges beneath them.[182] Nevertheless, the utilization of district court data requires significant adjustment to a basic agnostic model.[183] In addition to those issues, there are several other factors that must be included in a valid model of ideology of U.S. courts of appeals judges.[184] Each of those concerns is discussed below.[185]

A. *Data Gathering*

For this study, data was gathered from opinions issued by the First, Second, Third, Fourth, Fifth, Sixth, Seventh, Eighth, Ninth, Tenth, and

---

[176] *See* Choi & Gulati, *Tournament, supra* note 108, at 303 ("The selection of future Supreme Court justices on the basis of such objective criteria would make clear (and thereby reduce) the role that politics plays in both the initial process of selecting a candidate and the often highly political Senate confirmation proceedings.").

[177] *See id.* ("The norm today appears to be that a candidate for the Supreme Court must first sit on a federal circuit court of appeals before [he or] she may be considered for a seat on the court."); O'Neill, *supra* note 26, at 702.

[178] *See* Fischman & Law, *supra* note 7, at 162–63.

[179] *See id.*

[180] *See id.* at 165.

[181] *See infra* notes 186–243 and accompanying text.

[182] *See infra* notes 186–191 and accompanying text.

[183] *See infra* notes 200–242 and accompanying text.

[184] *See infra* notes 282–314 and accompanying text.

[185] *See infra* notes 186–314 and accompanying text.

Eleventh Circuit Courts of Appeals in 2008. The resultant data set ("Case Database") from those circuits included 30,726 judicial votes from panel decisions. The Case Database included opinions that used a standard of review while excluding immigration[186] and habeas corpus[187] cases due to the unique standard of review issues in such cases.[188] In an effort to address a substantial defect in the major data sets that have been used to empirically study the U.S. courts of appeals, the Case Database included unpublished opinions available through electronic databases.[189] The Case Database also included the district court decisions underlying the appellate review in every instance where a district court judgment was reviewed.

---

[186] *See* Hiroshi Motomura, *Immigration Law and Federal Court Jurisdiction Through the Lens of Habeas Corpus*, 91 Cornell L. Rev. 459, 474 (2006) (discussing how large portions of immigration case review are based upon a collateral review model that affords a very high level of deference because "[t]he regulations . . . revise the standard of review to require greater deference to an immigration judge's findings of fact").

[187] *See* Brandon Scott, *When Child Abuse Becomes Child Homicide: The Case of* Gilson v. Sirmons, 34 Okla. City U. L. Rev. 281, 293, 305 (2009) (discussing the "unique" standard of review in federal habeas cases due to the Antiterrorism and Effective Death Penalty Act).

[188] For each of the courts of appeals databases in LexisNexis, the following search was executed, and all of the results were downloaded and coded: "date aft 1/1/2008 and date bef 1/1/2009 and ("De Novo" or Clear! Erro! or (Arbitrar! w/3 Capricious!) or (Abus! w/3 Discretion) or "Substantial Evidence" or "Standard of Review") and not immigration and not habeas."

[189] *See* Edwards & Livermore, *supra* note 24, at 1923. According to Harry T. Edwards and Michael A. Livermore:

> [T]he Songer database does not include unpublished decisions issued by the courts of appeals. This is an extremely important omission, because a huge percentage of courts of appeals decisions are reported but unpublished. In fact, in 2007, less than 17 percent of all opinions in the courts of appeals were published. Published decisions as a sample of total decisions are far from random: the judgments rendered in unpublished decisions are largely unanimous, and these cases typically involve more straightforward applications of law. Unpublished decisions, no less than published decisions, dispose of appeals on the merits. Importantly, unpublished decisions offer valuable information regarding a court's adherence to precedent, because in these cases the law is often most clear. Law professors and researchers tend to focus on published decisions that raise difficult issues and establish new precedent. For the vast majority of litigants, however, it is often of no moment whether a case is published or not. The court's judgment is what matters. And every judgment counts when one attempts to accurately measure the work of the appellate courts. Therefore, any assessment of the work of the courts of appeals that does not include unpublished decisions cannot be seen as complete.

Id. (emphasis omitted).

Cases were coded for, among other variables: judges; whether individual judges were sitting by designation; disposition; type of case (e.g., criminal/environmental); prevailing party; circuit; district court judge; district court; whether the case involved constitutional review of legislation; whether the case was reviewing the decision of an executive agency; and the standard of review used. In analyzing each case, the vote of each judge on the panel was coded separately. This allowed for each judge's ideology to be determined independently, even when he or she dissented in a case.

In addition to the 2008 database of cases, a separate database ("Judge Database") was constructed that included biographical and demographic data about individual judges. In the Judge Database, judges were coded for, among other variables: appointing president; presidential party; rating by the American Bar Association at the time of nomination; age at the time of appointment; age as of 2008; composition of the Senate at the time of the confirmation vote; gender; race; law school attended; prior work experience; whether the president and majority of the Senate were of the same party at the time of the appointment; and whether the judge took senior status during or before 2008. The Judge Database featured data for all federal appellate judges who served on panels and were included in the Case Database, as well as district judges who issued opinions reviewed by those appellate courts.

This study's data set included data from 1293 judges who served on the U.S. courts of appeals, including judges who sat by designation and who served in the federal district courts. Most of those judges, however, only issued a few judicial votes. This was particularly true of judges who had taken senior status before or during 2008, district court judges who sat upon appellate panels by designation, or district court judges who only had a limited number of cases reviewed by the appellate courts.[190] Because the smaller sample sizes of votes by judges might offer less valid data, the discussion below often limits the analysis to the results from the 177 judges on the U.S. courts of appeals who had a significant number of interactions in the range of categories of cases studied.[191]

---

[190] In all, the data set included 2273 judicial votes by judges sitting by designation.

[191] The cutoff was conservatively based upon a relatively even distribution of judges with lower vote counts throughout the political spectrum.

B. *Determining the Ideology Score*

Although the core element of the metric is determining agreements and disagreements, a simple count of such instances is inadequate to effectively gauge ideology.[192] Judges are not situated in identical manners.[193] Most importantly, they have differing case type mixes, review district courts with varied degrees of deference, sit on different circuits, and interact with very dissimilar groups of judges.[194] As a result, the model incorporates each of those differences in an effort to capture a judge's ideology by including the context in which his or her judicial votes were made.

1. Adjusting for Case Mixes

Because each circuit is defined based upon geography, every circuit has a different caseload.[195] Further, judges within those circuits hear only a portion of those cases, which can create further inconsistencies among the types of cases each of the judges hears.[196] Because this measure relies on agreements and disagreements among judges, it is important to control for areas of case law where agreements are more or less likely.[197] This can be done by independently analyzing the data from each significant area of law for each judge. Doing so, however, can increase dramatically standard errors for the metric because the sample sizes within each sub-category of case type will be far less than if all of the judge's decisions were analyzed together.[198] Thus, there is a need to balance case type controls with the desire for adequate samples to draw valid statistical inferences.

In this data set, one major distinction in case types proved important in the commonness of disagreements among judges: criminal and civil cases. The level of disagreement was particularly pronounced between appellate and district judges. In criminal cases, appellate courts only reversed district court judgments in 14.6% of the cases. In contrast, over 26% of the judgments in civil matters were overturned by the

---

[192] *See* KUERSTEN & SONGER, *supra* note 23, at 2–4.

[193] *See id.*

[194] *See id.*

[195] *See* Paul R. Michel, *Assuring Consistency and Uniformity of Precedent and Legal Doctrine in the Areas of Subject Matter Jurisdiction Entrusted Exclusively to the U.S. Courts of Appeals for the Federal Circuit: A View from the Top*, 58 AM. U. L. REV. 699, 702 (2009) (describing the Federal Circuit as the only court of appeals that is based on subject-matter, not geography).

[196] *See* KUERSTEN & SONGER, *supra* note 23, at 2–4.

[197] *See id.*

[198] *See* LAWLESS ET AL., *supra* note 95, at 425; Fischman & Law, *supra* note 7, at 157–58.

U.S. courts of appeals. At each stage of the process, judges had their Ideology Scores determined as if they had been on panels with an average criminal/civil case mix to address the very different reversal rates. As a result, an individual judge's data included a numerical breakdown of the civil and criminal cases heard, as described in further detail in Section C.[199]

2. Incorporating Standards of Review

By introducing the district court into the model, certain difficulties emerged. Notably, the process of review at each court level is quite different.[200] The degree of deference between judges varies based upon the particular standards used and the hierarchal relationship of the judges being studied.[201] These standards of review could increase or decrease the likelihood of reversal of district court judgments and thereby affects a judge's Ideology Score.[202] This study, thus, incorporates appellate standards of review as part of including district court judgments in the model.

Standards of review are formal rules used by appellate courts to determine the degree of deference that they should give to lower court judgments.[203] These standards are regularly used in cases and are normally identified by judges in their opinions.[204] Standards of review do not directly dictate the outcome in a case. Judges are free to reverse a district court judgment when using a deferential standard and to affirm a district court using non-deferential review.[205] Consequently, if standards of review are serving a function, it is essential to incorporate the standard of review used by the appellate court for purposes of gauging judicial ideology using a multi-court measure.[206] For example, a decision by an appellate judge to reverse a conservative district court judge is more significant if a deferential standard of review was applied instead of a non-deferential one, such as de novo.[207] In the aggregate, an

---

[199] *See supra* notes 243–248 and accompanying text.

[200] *See* Cross, *supra* note 72, at 1500–02.

[201] *See id.*

[202] *See id.*

[203] *See id.*

[204] *See id.*

[205] *See id.*

[206] *See* Cross, *supra* note 9, at 228 ("[J]ust one legal standard, affirmance deference to the lower court decision, is consistently significant statistically and by far the most important single variable substantively in explaining circuit court outcomes."); Cross, *supra* note 72, at 1500–02.

[207] *See* Cross, *supra* note 9, at 228; Cross, *supra* note 72, at 1500–02.

appellate judge who reverses district court judges of a particular ideology when deferential standards are applied is substantially more extreme in his or her ideology than a judge who only reverses in cases using a de novo standard.[208] Similarly, a decision by an appellate court judge to affirm a liberal district court judge is less noteworthy when a deferential standard is applied because the affirmance might simply be a product of the standard of review.[209]

Existing empirical scholarship supports the notion that standards of review create meaningful differences in reversal rates.[210] A prior study by Frank Cross based upon the Songer Database for the U.S. courts of appeals found that standards of review are correlated with a change in reversal rates.[211] Nevertheless, because the Songer Database only coded standards of review in cases involving executive agency review, it is unclear if the results would extend to general application of such standards.[212] Based upon the results of the data set used in this Article, the application of a deferential versus a non-deferential standard of review is strongly correlated to a change in reversal rates in different types of cases.[213]

Figure 1, below, illustrates the results from all circuits in this study of the reversal rates with the three most commonly applied standards of review. The only non-deferential standard in Figure 1 is de novo review.[214] The other two standards afford deference to the judgments of district courts.[215] The results, which show a higher reversal rate for non-deferential standards than for deferential standards, are expected if standards of review serve their proper formal role by constraining appellate judges, such that those judges defer more often when standards of review dictate that they should do so.[216]

---

[208] *See* Cross, *supra* note 72, at 1500–02.

[209] *See id.*

[210] *See id.* at 1502–03.

[211] *See id.*

[212] *See id.* (coding only 808 cases).

[213] *See* Edward K. Cheng & Albert H. Yoon, *Does* Frye *or* Daubert *Matter? A Study of Scientific Admissibility Standards*, 91 Va. L. Rev. 417, 497 n.58 (2005). Conventionally, the value of p indicates a statistically significant relationship if it is less than .050. *See id.* P-values represent the probability that the results obtained were not due solely to chance. So, a p-value of .050 means that the results were 95% likely to be the result of something other than simple chance.

[214] *See* Cross, *supra* note 72, at 1502.

[215] *See id.*

[216] *See id.* at 1501–02. Nevertheless, if there were perfect knowledge and rational decision-making among the litigants, at least in private actions, there should be no observable difference in reversal rates because the parties would act accordingly. *See id.*

Figure 1 - Reversal Rate by Standard of Review (p=0.000)
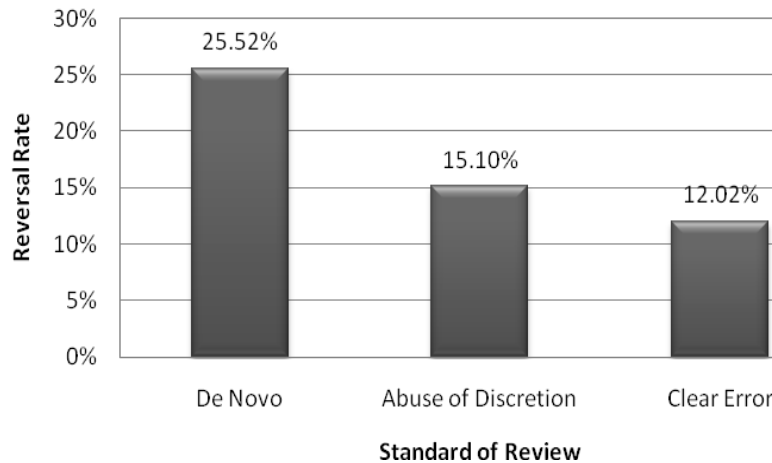
Given the significant difference in reversal rates between the non-deferential and deferential standards, incorporating reversal rates of different judges using deferential standards is a necessary step toward creating a valid multi-court measure of judicial ideology.[217] As a result, the judicial votes were divided into cases with deferential, non-deferential, and unclear standards of review.

3. Inter-Circuit Adjustments

Every federal Court of Appeals has, among other differences, varied substantive law, procedural law, appellate judges, and district court

---

[217] *See* CROSS, *supra* note 9, at 282; Edwards & Livermore, *supra* note 24, at 1908. The substantial role of standards of review is a significant rebuttal against the strong versions of the attitudinal and strategic models in regard to federal appellate court judges. *See* CROSS, *supra* note 9, at 288. The attitudinal model in particular, which contends that judges make decisions based primarily upon policy preferences, has served as the basis for many empirical studies. *See id.* at 11 ("[I]n '[other] corners of the university . . . it is widely considered a settled social scientific fact that law has almost no influence on the justices.'") (quoting Howard Gillman, *What's Law Got to Do With It? Judicial Behavioralists Test the "Legal Model" of Judicial Decision Making*, 26 LAW & SOC. INQUIRY 465, 466 (2001) (reviewing HAROLD J. SPAETH & JEFFREY A. SEGAL, MAJORITY RULE OR MINORITY WILL: ADHERENCE TO PRECEDENT ON THE U.S. SUPREME COURT (1999))); Edwards & Livermore, *supra* note 24, at 1908 ("[E]mpirical studies of judicial decisionmaking are too often informed by the 'attitudinal model' of judicial behavior."). Thus, under such a model, legal rules like standards of review that apply do not serve to constrain judicial decision-making. *See* CROSS, *supra* note 9, at 11–12. Yet the result illustrated above indicates that at least some legal rules like standards of review have a powerful effect on the outcome in cases.

judges.[218] Those differences raise the risk that, even if the measure successfully describes ideology within each circuit, it would still fail to offer valid comparisons between circuits.[219] Consequently, there is a concern any time a researcher compares the different U.S. courts of appeals that there might be unobserved variables that explain variation in results among the circuits.[220]

For example, if a circuit court of appeals was largely appointed by Republican presidents, but the districts within the circuit were filled with appointees of Democratic presidents, one might expect differences in deference when compared to a circuit with balanced appointees.[221] Similarly, differences in substantive law might explain why standards of review are applied in an uneven manner.[222] For this reason, the findings in regards to different levels of ideology among circuits, as well as inter-circuit comparisons of individual judges, would be called into question.[223]

This "baseline" issue emerges in this study in two ways.[224] First, each circuit likely has different concentrations of liberals and conservatives based upon the number of slots a particular appointing president filled.[225] For example, of the eleven active judges sitting on the Eighth Circuit, eight were appointed by President George W. Bush,[226] two by President Bill Clinton,[227] and one by President Ronald Reagan.[228] Assuming that the three presidents did a reasonable job in appointing

---

[218] *See* KUERSTEN & SONGER, *supra* note 23, at 2–4. With the exception of the Federal Circuit, which is not included in this study, the circuits have geographic and not subject matter based jurisdiction. *See* Michel, *supra* note 195, at 702.

[219] *See* KUERSTEN & SONGER, *supra* note 23, at 2–4.

[220] *See* David C. Vladeck, *Keeping Score: The Utility of Empirical Measurements in Judicial Selection*, 32 FLA. ST. U. L. REV. 1415, 1433–34 (2005) (discussing, in the context of Choi & Gulati, *Tournament, supra* note 108, the need to account for differences among circuit caseloads in creating empirical measures).

[221] *See* Cross, *supra* note 72, at 1503–04.

[222] *See id.*

[223] *See id.*

[224] *See generally* Brian Z. Tamanaha, *Devising Rule of Law Baselines: The Next Step in Quantitative Studies of Judging*, THE LEGAL WORKSHOP (Mar. 25, 2010), http://legal workshop.org/2010/03/25/2667 (submitted by DUKE L.J.).

[225] *See* Fischman & Law, *supra* note 7, at 149.

[226] *Eighth Circuit Court of Appeals Judges*, U.S. COURT OF APPEALS FOR THE EIGHTH CIRCUIT, http://www.ca8.uscourts.gov/newcoa/judge.htm (last visited Aug. 15, 2010). Chief Judge James B. Loken and Judges William J. Riley, Michael J. Melloy, Lavenski R. Smith, Steven M. Colloton, Raymond Gruender, William Duane Benton, and Bobby E. Shepard were all appointed by President George W. Bush. *Id.*

[227] *Id.* Judges Diana Murphy and Kermit E. Bye were appointed by President Bill Clinton. *Id.*

[228] *Id.* Judge Roger L. Wollman was appointed by President Ronald Reagan. *Id.*

judges with ideologies similar to their own, the Circuit would appear to be very conservative. Judges Diana Murphy and Kermit Bye, the two Clinton appointees on the circuit, sat on panels on which 90% of their co-panelists were appointed by Republican presidents. Further, because the rate of dissent in the data set is approximately 2%, it is quite likely that Judges Murphy and Bye would agree with far more Republican appointees than Democratic ones. That Judges Murphy and Bye agreed with Republicans more often would be expected simply due to random panel assignment and not their actual ideologies. Thus, a valid model must account not just for the actual levels of agreement and disagreement, but also the expected levels of agreement and disagreement.[229] The effect is also duplicated among circuits with different political concentrations of district court judges.[230]

Second, at least some body of cases in the data set was decided with no regard to the judge's ideology.[231] This is reflected by the high level of agreement among appellate judges[232] as well as prior studies of the U.S. courts of appeals.[233] For this reason, it can be very difficult to be sure that the decisions are not the result of some unobserved variable such as clear precedent or strategic, personal goals of a judge.[234] These unobserved variables might be different in each Circuit.

The measure accounts for these difficulties at a general level by relying on rates, and not frequencies, of agreement. So, if a judge sat with twice as many judges appointed by Republican presidents, the measure would not be distorted because the percentage of agreement in the examined category is what is used in the final computations. Further, each category of cases (i.e., criminal or civil) is weighted in the final score based upon the rate at which the average judge in the data set reviewed such cases. Based upon the political composition of the panel (including the district judge), standard of review used, and type of case, an expected score was computed for each judge.

Further, there are two specific reasons to believe that the inter-circuit differences relevant to this study are effectively accounted for. First, the similarity of reversal rates among the various circuits weighs strongly in favor of a proposition that the differences among the cir-

---

[229] *See* Fischman & Law, *supra* note 7, at 162–63.

[230] *See id.*

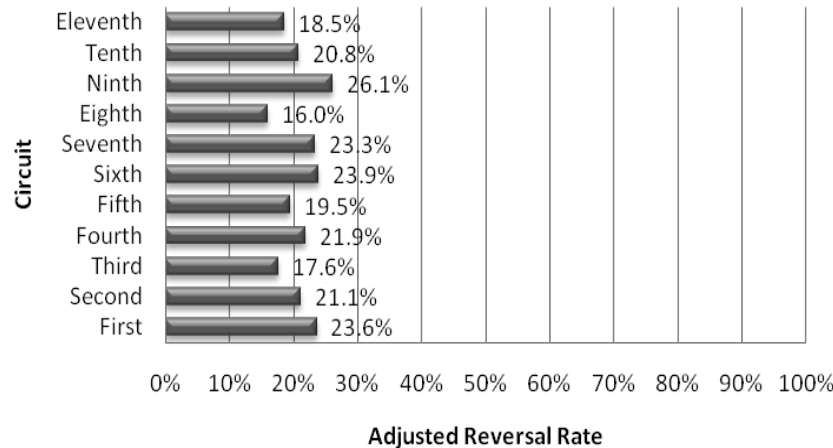[231] *See* Fischman, *supra* note 162, at 3–4.

[232] In the data set, the overall agreement rate among judges on the Courts of Appeals in all cases was 99.4%.

[233] *See* Fischman, *supra* note 162, at 3–4.

[234] *See* POSNER, *supra* note 59, at 29.

cuits are not as large as might be contended. An "Adjusted Reversal Rate," which equalizes the number of criminal and civil cases in the Circuit data, was derived for each circuit. As Figure 2 illustrates, the differences among the circuits were relatively small.

Figure 2 - Adjusted Reversal Rate by Circuit

Circuit	Adjusted Reversal Rate
Eleventh	18.5%
Tenth	20.8%
Ninth	26.1%
Eighth	16.0%
Seventh	23.3%
Sixth	23.9%
Fifth	19.5%
Fourth	21.9%
Third	17.6%
Second	21.1%
First	23.6%

The circuit least likely to reverse was the Eighth, which reversed at a rate of 16.0%. The circuit most likely to reverse was the Ninth with an Adjusted Reversal Rate of 26.1%. If political differences between the appellate and district court judges in a circuit were a significant factor driving judicial votes, one would expect the overall reversal rate to have been higher relative to other circuits. Similarly, variations among substantive law would have been expected to surface in the overall reversal rates.

Second, the circuits are not entirely insular.[235] Senior judges often travel between circuits and issue opinions in each location.[236] These points of interaction between the judges on different circuits can provide bridges for comparison among judges regardless of their home circuit.[237] In this data set, there were 2472 votes by twenty-six judges who sat on panels in more than one circuit.[238] In every instance where

---

[235] *See* Kelly Baker, Note, *Senior Judges: Valuable Resources, Partisan Strategists, or Self-Interest Maximizers*, 16 J.L. & POL. 139, 150 (2000) ("Senior judges may travel among circuits and districts to provide services as necessary.").

[236] *Id.*

[237] *See id.*

[238] The traveling judges included in the data set were Judges Arthur L. Alarcon, Ruggero J. Aldisert, Bobby R. Baldock, Clarence A. Beam, Pasco M. Bowman II, Myron H.

one of the twenty-six judges voted, their Ideology Scores were broken down by the circuit in which the vote was registered. At the end, the differences in Ideology Scores for the collective group of travelling judges were calculated. The differences were further transformed into a per judge measure. These per judge measures were applied to each judge in the circuit in computing the Ideology Scores for those judges.

4. Panel Effects

Consistent with prior studies of federal appellate courts, so-called "panel effects" were included in the measure.[239] Panel effects are the degree to which the ideologies of a judge's co-panelists change that judge's vote.[240] Panel effects are significant for the Ideology Scores insofar as those Scores depend upon panel agreement rates and panel effects alter those agreement rates. To provide the best measure possible, the Ideology Scores should be adjusted to determine values as if a judge experienced no panel effects.

Although prior research on panel effects has focused exclusively on appellate co-panelists, this study assumes that district court judges are hidden fourth members of the appellate panel that form a larger "super panel." Notably, in a significant development beyond prior research, the data indicates that the party of the appointing president of the district judge, in relation to the panel judges, is correlated with the voting of the panelists reviewing the district court judgment.[241]


Bright, Robert E. Cowen, Richard D. Cudahy, David M. Ebel, Joseph J. Farris, William L. Garwood, John R. Gibson, Neil M. Gorsuch, David R. Hansen, Paul J. Kelly Jr., Robert B. King, Gilbert S. Merritt Jr., Robert J. Miner, Karen N. Moore, Jon O. Newman, Jane R. Roth, Eugene E. Siler Jr., Walter K. Stapleton, A. Wallace Tashima, John M. Walker Jr., and J. Clifford Wallace.

[239] *See* Fischman & Law, *supra* note 7, at 149–50 ("One challenge that empirical scholars must address, therefore, is the fact that panel composition effects can conceal the true extent of a judge's ideological leanings. Because the influence of ideology on a judge's voting behavior may be muted unless he or she is paired with at least one likeminded colleague, a simple analysis of individual judicial voting records that fails to control for panel composition is likely to underestimate the true extent of the judge's ideological preferences.").

[240] *See* Cross & Tiller, *supra* note 89, at 2173–75 (finding that Republican and Democrat-controlled panels were more likely to defer to agency decisions that corresponded with their majority's ideology); Revesz, *supra* note 89, at 1768 ("[T]o a surprisingly strong extent, a judge's vote is affected by the identity of her colleagues on the panel.").

[241] Although a brief discussion of the panel effects findings are included here to provide the best possible Ideology Scores, the exact details of those findings are beyond the scope of this Article. The panel effects issues raised by the incorporation of the district court judge warrant a separate piece.

To understand the relevance of this correlation, it is helpful to outline the various scenarios in which the relevant judges might be arranged. There are three basic alignments of panels and district judges, based upon the party of the appointing president: (1) the two co-panelists and district judge might be of the same party; (2) the two-panelists might be of the same party and the district judge is of the opposite party; and (3) the co-panelists are of different parties and the district judge is of the same party as one of the co-panelists. The varying disagreement rates on the panels in each of those situations are illustrated in Figure 3.

Figure 3 - Panel Disagreement Rate by Appointing President's Party of Co-Panelists and District Judge
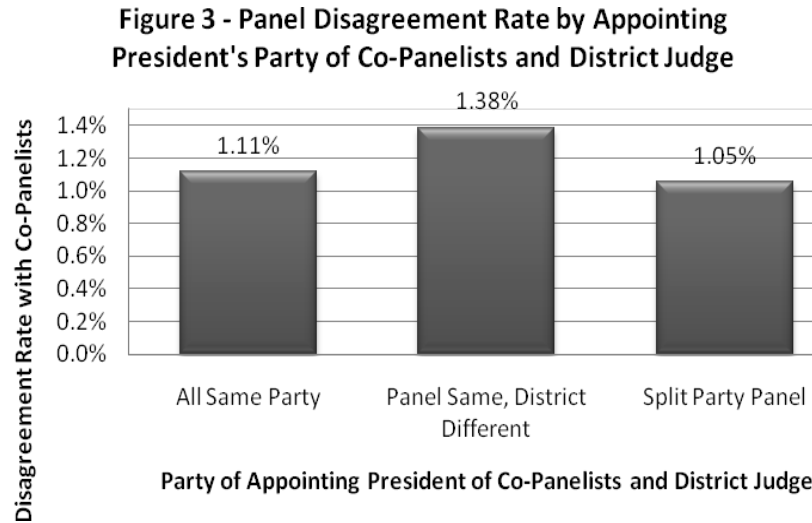

Although the absolute levels of disagreement (all below 1.50%) appear small, the ratio between them is significant. If the co-panelists were of the same party, but a different party from the district judge, the likelihood of disagreement among the studied judge and his or her co-panelist increased substantially. Such situations created a disagreement rate almost 25% higher than when the panel judges were appointed by presidents of differing parties and over 31% greater than situations where the co-panelists and district judge were all of the same party. The results indicate that the "super panel" alignment has effects on panel judge agreement and must be included in the scoring system.

Additionally, the political alignment among the co-panelist appellate judges and district judge affects a judge's decision to affirm or reverse the judgment of the district judge. Figure 4 illustrates the differences in reversal rates based upon the three alignments described above. Based upon their appointing president, if the two co-panelists

are of the same political party, but a different party than the district judge, then a judge is approximately 11% more likely to reverse the judgment of the district judge than in either of the two possible alignments.

Figure 4 - Reversal Rate by Appointing President's Party of Co-Panelists and District Judge

Party of Appointing President of Co-Panelists and District Judge
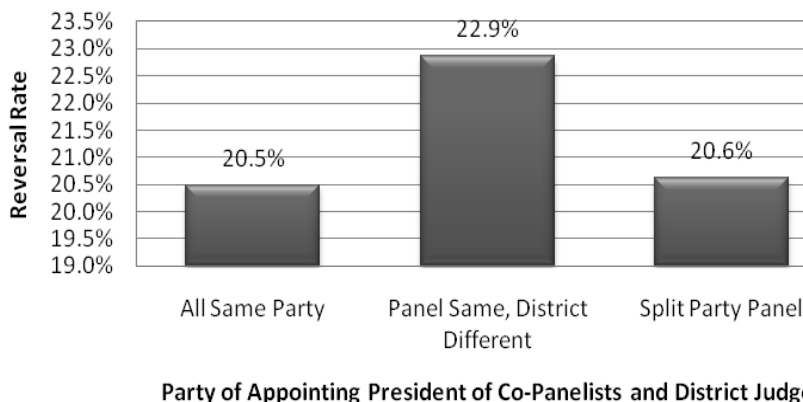
The three major alignments can be disaggregated into six possible situations based upon the specific ideologies involved. Listing the political parties of the two co-panelists followed by the district judge party (based upon appointing president), those six alignments are: (1) DD-D; (2) DD-R; (3) DR-D; (4) DR-R; (5) RR-D; and (6) RR-R. The significance of the "super panel effects" described above for the Ideology Scores is based upon the rate at which a judge is situated in each of the six scenarios. If a judge encounters more DD-R than RR-D situations, it would be expected that the studied judge's agreement rate with judges appointed by Democrats would increase over what would be expected with no panel effects. Also, that judge would be expected to reverse the judgments of more Republican appointees. Thus, a measure of ideology that incorporates the district court should integrate the rate at which a judge encounters the possible political alignments. The specific application of the methodology for adjusting the Scores based upon those panel effects is described below in Section C.[242]

---

[242] *See infra* notes 243–248 and accompanying text.

C. *The Multi-Court Agnostic Approach and Its Advantages*

Based upon all of the above considerations, the Ideology Score includes several components. Initially, there is the Panel Score (PS), which is a judge's agreement level with co-panelists. Similarly, a District Score (DS) is based upon agreements (affirmances) with the district judge being reviewed. An agreement represents a vote to affirm, and a disagreement represents a vote to reverse, the lower court's judgment. An agreement with a liberal judge is included as a negative value, and with a conservative judge, a positive value in order to best represent the common depiction of the left-right political spectrum. For there to be a valid agnostic model, there needs to be prior assessments of some judges. In this model, every judge is initially assigned a score based upon the ideology of the appointing president (-1 for a Democratic appointee and 1 for a Republican one). For the judge who is having an Ideology Score determined, however, he or she is treated as a blank slate with no prior indication of ideology. The Panel Score expression is as follows:[243]

> *Panel Score = AvgCrim * (RateAg Rep – RateAg Dem) + AvgCiv * (RateAg Rep – Rate Ag Dem)*

The District Score was more complicated because of the need to incorporate standards of review. It was determined as follows:[244]

> *District Score = (AvgCrim * (WeightDef Crim * NumDef * (RateAff Rep – RateAff Dem) + WeightNon Crim * NumNon * (RateAff Rep – RateAff Dem) + WeightOth Crim * NumOth * (RateAff Rep – RateAff Dem)) + AvgCiv * (WeightDef Civ * NumDef * (RateAff Rep – RateAff Dem) + WeightNon Civ * NumNon * (RateAff Rep – RateAff Dem) + WeightOth Civ * NumOth * (RateAff Rep – RateAff Dem)))/(NumDef + NumNon + NumOth)*

In addition to the Panel Score and District Score, there was a "Panel Effects Adjustment" made based upon the observed panel effects in the data. The equation for the raw total scores is as follows:

---

[243] AvgCrim = Average percentage of criminal cases; AvgCiv = Average percentage of civil cases; RateAg = Rate of agreement; Rep = Judges appointed by Republican presidents; Dem = Judges appointed by Democratic presidents.

[244] RateDef = Rate of agreement with deferential standard of review; RateNon = Rate of agreement with non-deferential standard of review; RateOth = Rate of agreement with unclear standard of review; WeightDef = Weight applied to deferential cases based upon difference in reversal rates; WeightNon = Weight applied to non-deferential cases; WeightOth = Weight applied to cases with an unclear standard of review.

Raw Score = Panel Score + District Score – Panel Effects Adjustment

For example, the raw Ideology Score for Judge J. Harvie Wilkinson III was calculated as follows. In regards to his co-panelists on the Fourth Circuit, his agreements are illustrated in Figure 5 below. As the rate of agreement was very high for the federal appellate courts in the data, Judge Wilkinson only had five disagreements, two with Democratic appointees in civil cases, one with a Democratic appointee in a criminal case, and two with Republicans in civil cases.

Figure 5 - Judge Wilkinson's Panel Score Components
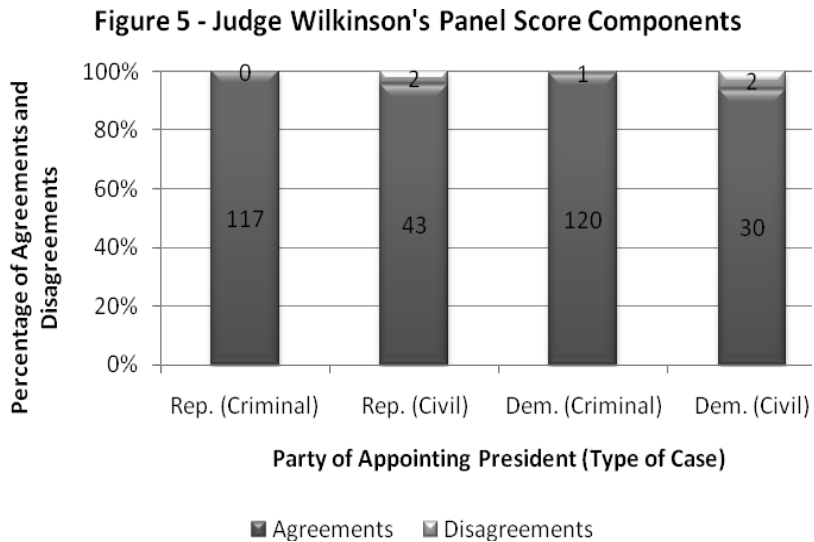
So, using the above equation, Judge Wilkinson's Panel Score was determined as follows: 0.442 (the average percentage of criminal cases) * (117/117 – 120/121) + 0.558 (the average percentage of civil cases) * (43/45 – 30/32), totaling 0.01373. In reviewing the judgments of district courts, Figure 6 illustrates the breakdowns of Judge Wilkinson's agreements with those courts based on affirmance rates in cases with different standards of review. Notably, the rate of disagreement (reversals) was much higher with district court judges than with judges on the courts of appeals.

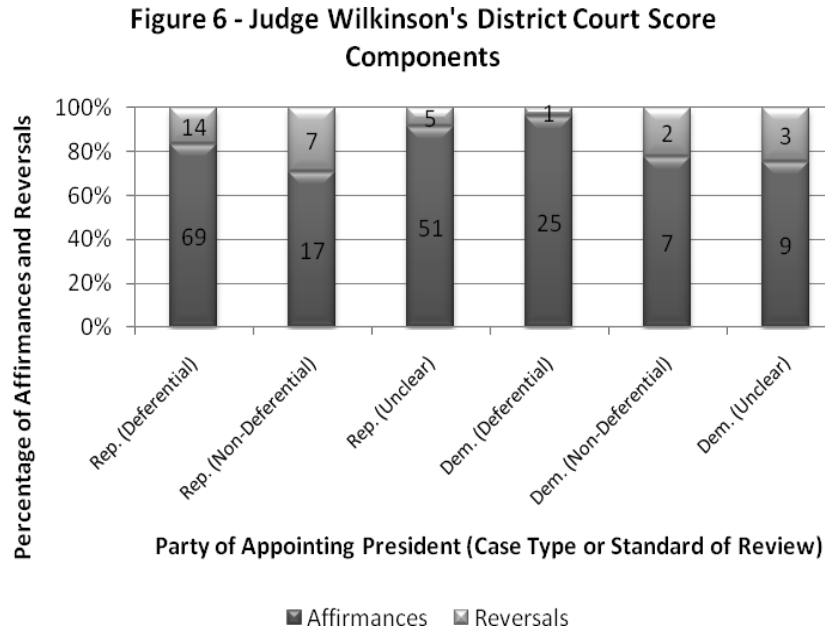Figure 6 - Judge Wilkinson's District Court Score Components

Figure 6 does not include the breakdown of cases based upon review of civil or criminal matters, but the allocation of those case types was incorporated into the Ideology Scores as follows. Judge Wilkinson's District Score was: (0.446 (the average percentage of criminal cases)[245] * (0.479 (weight for criminal cases with a deferential standard)[246] * 119 * (69/83-25/26) + 0.241 (weight for criminal cases with a non-deferential standard) * 33 * (17/24 – 7/9) + 0.280 (weight for criminal cases with an unclear standard) * 68 * (51/56 – 9/12)) + 0.554 (the average percentage of civil cases) * (0.384 (weight for civil cases with a deferential standard) * 119 * (69/83-25/26) + 0.328 (weight for civil cases with a non-deferential standard) * 33 * (17/24 – 7/9) + 0.289 (weight for civil cases with an unclear standard) * 68 * (51/56 – 9/12))) / (119 + 33 + 68). Because the respective weights added up to two, the total was divided in half to keep the District Score on the same relative scale (0 to 100%) as the other components. This yielded a District Score of -0.0110.

---

[245] Because data was incomplete for some district or appellate judges, the average percentages of civil and criminal cases for the Panel and District Scores vary slightly.

[246] The weightings were based upon the differential in reversal rates in each situation. So, deferential standard cases counted more than the non-deferential standard cases because reversal was rarer in such situations.

The last portion of the raw Ideology Score constituted the Panel Effects Adjustment. As mentioned above, the combination of the parties of the appointing presidents of a judge's co-panelists and the district judge under review had a statistically significant effect on a judge's vote. In order to determine Judge Wilkinson's Panel Effects Adjustment, it was essential to tabulate the differing alignments in which Judge Wilkinson participated. Figure 7 includes the relevant alignments based upon the party of the appointing president for the two co-panelists and the district judge under review.

Figure 7 - Judge Wilkinson's Panel Effects Adjustment Components
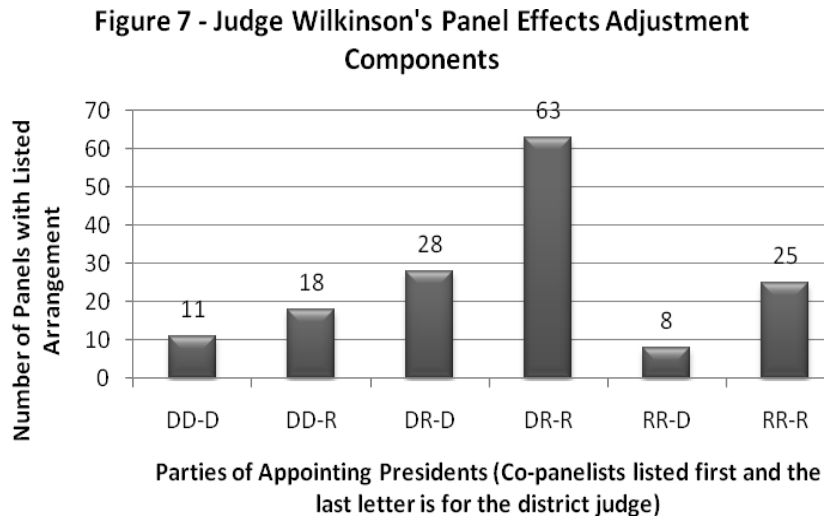
Each of the above six scenarios in Figure 7 were incorporated into the Panel Effects Adjustment. Of particular note, Judge Wilkinson's experience in 2008 indicated that he sat on a disproportionate number of panels reviewing district judges appointed by Republican presidents. Consistent with prior research, panel effects were assumed to be symmetrical in that co-panelists in either ideological direction exercise the same pull. Because the judge studied was treated as an ideological unknown, the key to panel effects were the other panelists and the district judge under review.

For each DD-D alignment, a judge was expected to agree at a rate of 98.89% as outlined in the previous Section. The same rate of agreement is expected in each RR-R scenario. Thus, each of those polar opposite scenarios canceled out the panel effect of the opposite scenario. Any excess in either direction was accounted for. Similarly, DD-R and RR-D situations canceled each other out and the excess number was

the critical factor for the Panel Effects Adjustment. Split panels with either a Democrat or Republican-appointed district judge also offset each other.

For Judge Wilkinson, there was an excess of 14 (25 – 11) RR-R over DD-D scenarios, 10 (18 – 8) DD-R over RR-D situations, and 35 (63 – 28) DR-R over DR-D alignments. The net result pulled Judge Wilkinson in a rightward direction compared to both his panel agreements and decisions to affirm or reverse. The percentage that each of the excess values represented of a judge's total votes provided the basis for the weighting of that panel effect. Based upon the average panel effects values described above, Judge Wilkinson's Panel Effects Adjustment was 0.000965. A similar "super panel" effects adjustment was made based upon the expected reversal rates in each of the above scenarios. That Panel Effects Adjustment was 0.002375. Combining the Panel Score, District Score, and Panel Effects Adjustment yielded a Raw Ideology Score of -0.0006169.

Three final adjustments were made to yield the final Ideology Score. First, the Scores were adjusted for the circuit on which the judge primarily sat, based upon the circuit differentials determined from the data of traveling judges as described in Section B.3.[247] The circuit differentials assumed that a judge's ideology would remain constant when traveling between circuits. So, if a judge exhibited a more liberal score in an "away" circuit over a "home" one, this implied that something about the "away" circuit (i.e., the circuit's substantive law or case mix) made judges more apt to have voted in a liberal direction. As a result, a circuit differential adjustment was subtracted from the judge to remove the effects of being in a particular circuit. For Judge Wilkinson, who sat on the Fourth Circuit, his Raw Ideology Score was adjusted from -0.0006169 to -0.0003216 because travelling judges voted more conservatively while in the Fourth Circuit.

Second, because the Scores were relative, it was necessary to identify a midpoint judge. Because 68 of the 177 judges were appointed by Democratic presidents, the zero value was assigned to the average of the 68th and 69th most liberal judges.

Third, to put the scale into values that were more meaningful, the Scores were scaled so that the possible range could extend from nega-

---

[247] *See supra* notes 218–238 and accompanying text. The adjustments for each of the circuits were determined to be as follows. First, -0.000332943; Second, -0.000258377; Third, 5.70766E-05; Fourth, -0.000307391; Fifth, -0.000750237; Sixth, 0.001157225; Seventh, 0.000560393; Eighth, -0.000159758; Ninth, -1.12073E-05; Tenth, 4.58607E-05; and Eleventh, -0.000865095.

tive 100 for the most liberal judge and positive 100 for the most conservative judge. Because the most liberal judge in the data set was further away from the judge with a 0 ideology, he had a resultant Ideology Score of -100. The most conservative judge had a score of 96.4. In the end, Judge Wilkinson had a score of 5.7, identifying him as a moderate conservative.

The Ideology Scores for a select portion of the judges are listed in the Appendixes to this Article. The distribution of the scores of the 177 judges with adequate sample sizes is indicated in Figure 8.

Figure 8 - Distribution of Ideology Scores (Mean = 7.5, Standard Deviation = 26.8)
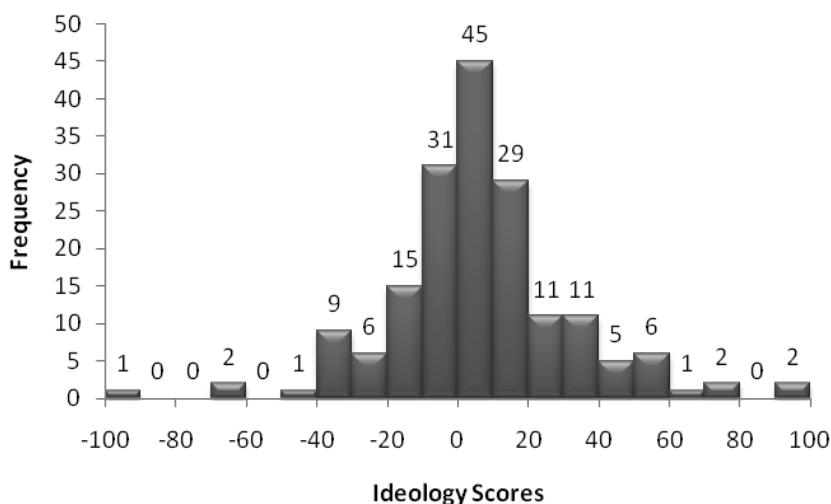
As would be expected, because Republican presidents have appointed the majority of judges on the federal appellate bench, there is a definite rightward distribution to the Ideology Scores of the 177 judges summarized in Figure 8. Beyond the right-leaning nature of the histogram, the distribution is generally normal, with a heavy concentration of judges near the zero point representing the point of an ideological moderate on the relative scale.

As an example of a distribution of scores among a specific circuit, Figure 9 includes the results from the Seventh Circuit along with Common Space Scores. To provide easier comparisons, the Common Space Scores were scaled from -100 to 100 in a manner identical to the technique used for scaling the Ideology Scores.

Figure 9 - Ideology and Common Space Scores of Seventh Circuit Judges
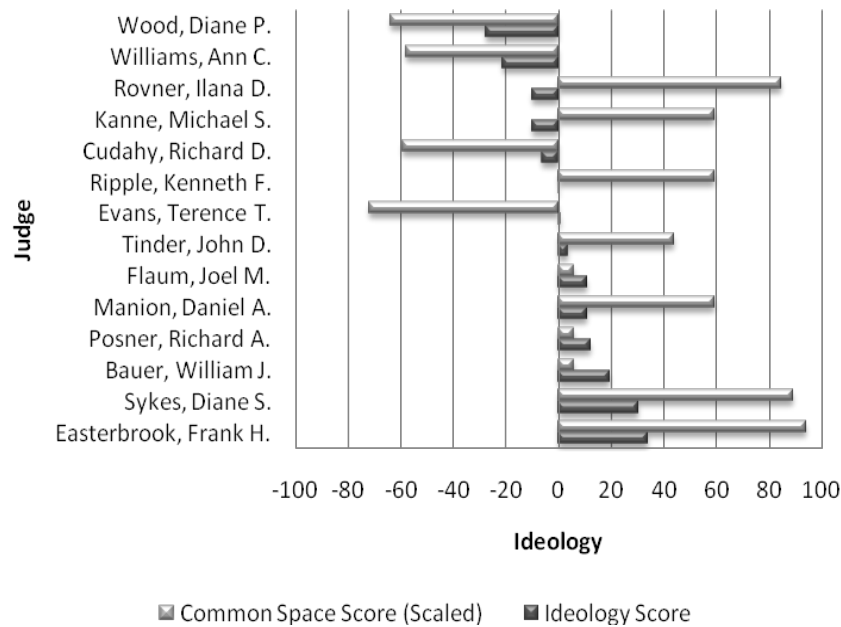
Although the judges in the Seventh Circuit largely follow the pattern of the most liberal judges (as determined by the Ideology Scores) appointed by Democratic presidents, there are some notable exceptions. Judge Terence Evans, for example, was found to be a slight conservative judge even though his Common Space Score would indicate he was the most liberal of all the Seventh Circuit judges. Judges Kanne and Rovner also had Scores pointing in opposite ideological directions. Although the Common Space and Ideology Scores agreed that Judges Sykes and Easterbrook were the most conservative, the magnitude of those differences was quite substantial. Overall, the Ideology Scores indicated the Seventh Circuit lacked a high degree of polarization, whereas the Common Space Scores showed substantial ideological variation.

In contrast with the Seventh Circuit, the Ideology Scores from the Sixth Circuit indicated a high degree of ideological polarization. Figure 10 includes the Ideology and Scaled Common Space Scores for Sixth Circuit judges.

Figure 10 - Ideology and Common Space Scores of Sixth Circuit Judges
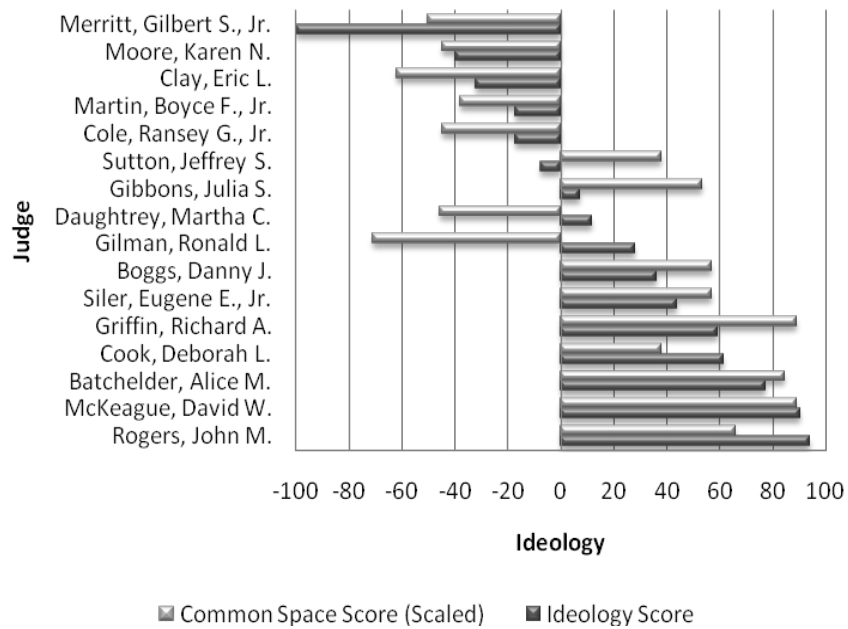
Even with the circuit differential adjustment, the Sixth Circuit had the most liberal judge in the data set (Judge Merritt) and the two most conservative (Judges Rogers and McKeague). The ideological battle lines seem much clearer in the Sixth Circuit than they were in the Seventh. Notably, there is a high level of consistency between the Common Space and Ideology Scores for the most ideological judges in the circuit. In contrast, the measures diverged sharply in the Eighth Circuit, as exhibited in Figure 11.

Figure 11 - Ideology and Common Space Scores of Eighth Circuit Judges

☑ Common Space Score (Scaled)   ■ Ideology Score
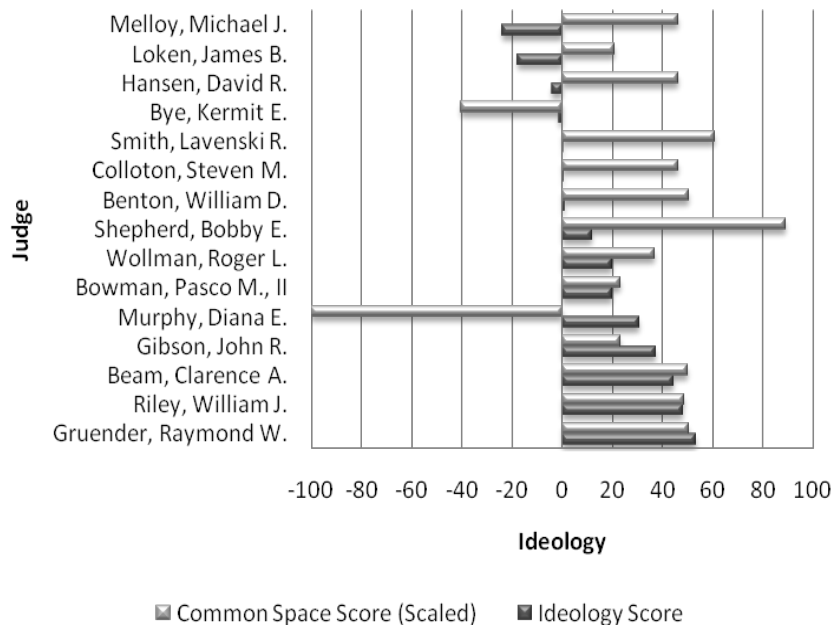
Whereas the Common Space Scores considered Judge Shephard to be the most conservative judge in the circuit, the Ideology Scores place him as only a slight conservative. Whereas the Common Space Scores identified Judge Murphy as the most liberal judge in the country, the Ideology Scores put her in the moderate conservative group. Indeed, many of the most liberal judges according to the Ideology Scores had among the most conservative Common Space Scores.

In each of the eleven circuits studied, there were numerous differences between the Ideology and Common Space Scores. Although the variations are worth noting, the larger question remains: which measure is "right?" In the next Part, the Article will address whether the assessments of the Ideology Scores are in any sense "better" than the existing measures.[248]

---

[248] *See infra* notes 249–281 and accompanying text.

III. Results and Discussion

Although the introduction of a new way to measure the ideology of judges on the U.S. courts of appeals is an important goal of this Article, there are two other questions that logically follow from such a scoring system. First, is this measure an improvement in a statistical sense over the alternatives that are available? Second, assuming there is scholarly value in the Ideology Scores, are they correlated with other significant components of a judge's background or demographics such that the judge's ideology could be predicted? Each of those questions is answered below.

A. *Ideology Predictions*

Assessing how well a particular statistical model of judicial behavior "works" is not an easy task.[249] Typically, models of this sort are assessed based upon how well they predict behavior.[250] Nevertheless, it is almost certainly unrealistic that the model would be able to predict all outcomes. What, then, constitutes success? In the area of judicial ideology measures, there has been very little scholarship concerning how various measures perform in practice.[251] Professors Joshua Fischman and David Law recently published the first major performance test of the ideology measures for federal judges below the Supreme Court.[252] They found little difference between the predictive power of the appointing president's political party and Common Space Scores in analyzing asylum claims in the Ninth Circuit.[253]

In an effort to determine how the two leading measures perform against the Ideology Scores, there needs to be data that is strongly indicative of ideology, but is otherwise unrelated to the measures being tested. As this Article notes at the outset, effective measures of the ideology of judges serving on the U.S. courts of appeals do not exist, so this is somewhat of a research paradox. If a truly valid and reliable baseline could be constructed, it would be the new leading measure of ideology instead of a statistical tool to assess other measures. Of course, how would a researcher know that this new measure was truly the best without a valid and reliable measure to compare it against?

---

[249] *See* Fischman & Law, *supra* note 7, at 172.

[250] *See id.* ("[P]racticing lawyers and social scientists want to know how judges *will in fact* behave.").

[251] *See id.* at 190–91.

[252] *Id.*

[253] *See id.* at 200–04.

This Article follows the lead of the recent study by Fischman and Law and applies the three measures of ideology to actual cases to see how they perform in predicting the outcomes. Because there is simply no other available data set concurrent with the timing and courts analyzed in this study, there was no viable alternative but to use this study's data set. In the Case Database, there are large numbers of cases in the civil rights area. In all, there were 1539 panels that considered claims under Title VII, 42 U.S.C. § 1983, and the Americans with Disabilities Act. It is generally assumed that a liberal judge will more often find in favor of the civil rights plaintiff. Consequently, each of the cases in those areas was coded based upon the prevailing party. The outcome was assumed to be liberal if a civil rights plaintiff won at the appellate level and conservative in the reverse circumstances.

There is, notably, a problem with testing the Ideology Scores in this study with the data set from which those scores are derived: circularity.[254] Circularity means that if the underlying data produces the model, the model should do well in predicting that data.[255] This would give the Ideology Scores an apparent strong "home-field" advantage.[256] To avoid this problem, the data used to assess predictions was removed from the computations of the underlying Ideology Scores. So, for the 1539 panels used to test predictions, the votes were excluded from the calculations that produced the modified Ideology Scores.[257]

Probit regressions were performed to determine if there was a correlation between the ideology metrics and the outcomes in the excluded civil rights cases using each of the three measures.[258] Of the 1539 cases, 907 had scores for all three measures and either a clear victory or loss for the plaintiff. Notably, the data indicated that the modified net and median Ideology Scores for the panels had statistically significant relationships with the outcomes in civil rights cases (p=0.00015 for net values and p=0.0002 for median values) with coefficients indi-

---

[254] *See id.* at 183 n.220. Circularity, also known as endogeniety, is the condition where the data being examined is also the source of the model being used. *See id.*

[255] *See* Fischman & Law, *supra* note 7, at 183 n.220.

[256] *See id.*

[257] As a result of the exclusion, there were no identified civil rights cases included in the computation of the modified Ideology Scores.

[258] *See* Lawless et al., *supra* note 95, at 414. Logistic or probit regression is used instead of linear regression when the outcome is categorical. *Id.* Because the case coding only afforded two options (criminal defendant win or loss), probit regression was used. *See id.*

cating that liberal judges were more likely to rule for plaintiffs.[259] Common Space Scores and the appointing president party did not perform as well. Net Common Space Scores did not have a statistically significant relationship (p=0.1419) with a civil rights plaintiff victory, but the median Common Space Score did (p=0.0473). Neither the net (p=0.2962) nor median (p=0.1306) appointing party value had a statistically significant relationship to the case outcome.

        Figure 12 illustrates that liberal Ideology Scores were correlated in a statistically significant manner (p=0.0007) with a civil rights plaintiff's winning percentage before the judge and Figure 13 illustrates a similar relationship for Common Space Scores (p=0.0078).

Figure 12 - Civil Rights Plaintiff Winning Percentage by Ideology Scores (p=0.0007)
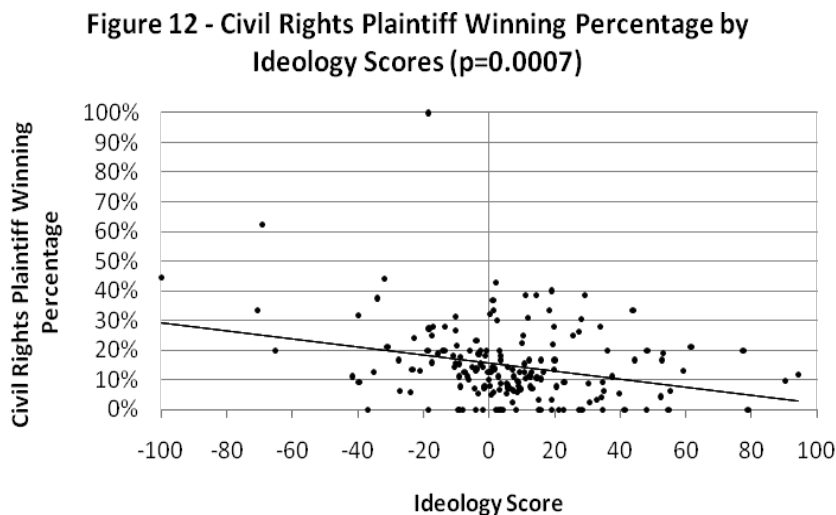
The downward slope of the line in Figure 12 indicates the degree to which being before a liberal judge (according to the Ideology Scores) increased the civil rights plaintiff's chance of winning. For roughly every ten Ideology Score points in a liberal direction, a judge is 1.4% more likely to vote for a civil rights plaintiff. Notably, the expected relationship between liberalism and civil rights victories for plaintiffs is far less clear with the Common Space Scores.

---

[259] For net Ideology Scores, the coefficient was -0.0147, and for the median Ideology Scores it was -0.0149. In either case, the direction of the coefficient supports the expected outcome that judges with liberal Ideology Scores would more often find in favor of civil rights plaintiffs.

Figure 13 - Civil Rights Plaintiff Winning Percentage by Common Space Scores (p=0.0078)
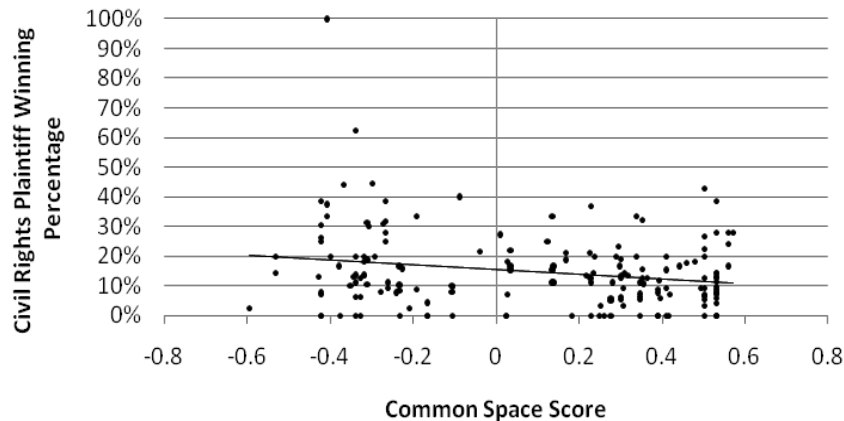
Although the statistical significance tests and results in Figures 12 and 13 are notable, further analysis provided greater confirmation of the predictive success of the Ideology Scores in civil rights cases. There is a variety of statistical tests to determine how well a model performs in predicting a particular set of outcomes. Chief among them is a method referred to as Pseudo R^2, which is meant to "measure[] both the model's goodness of fit and the strength of the association between the independent variable(s) and the dependent variable" using logistic and probit regression techniques.[260] In every test, the Ideology Scores had a higher Pseudo R^2 score than the Common Space Scores, which indicates a better "goodness of fit."[261] Even though it is difficult to deter-

---

[260] *See* Lawless et al., *supra* note 95, at 421.

[261] *See id.* at 411 (defining goodness of fit as, "in regression analysis, the concept of how well a regression model explains the data under examination"). Using the "fitstat" command in Stata, the following Pseudo R^2 results were recorded for the Ideology Scores and Common Space Scores. The first number is the value representing the net Ideology Score for the panel, the second is the value representing the median value for the Ideology Scores of the panel, the third is the net Common Space Score of the panel, and the fourth is the median Common Space Score for the panel. Stata Pseudo R^2: 0.0155, 0.0215, 0.0033, 0.0060; McFadden's R^2: 0.015, 0.021, 0.003, 0.006; Cragg-Uhler R^2: 0.022, 0.030, 0.005, 0.008; Efron's R^2: 0.015, 0.021, 0.003, 0.005; McKelvey & Zavoina's R^2: 0.031, 0.044, 0.006, 0.011. The overall low values of the Pseudo R^2 values even for the Ideology Scores was expected as a large majority of the criminal law decisions do not appear to be driven by ideology (as demonstrated by the high level of agreement). Consequently, the difference in scores simply illustrates the degree to which the Ideology Scores explain the variance based upon ideological decision making.

mine the degree to which the Ideology Scores outperformed the existing measures because of peculiarities associated with the computation of Pseudo R^2 values, the results indicate that such a difference likely exists.[262]

B. *Predicting Ideology*

Of significant value to politicians and scholars is identifying variables that are correlated with judicial ideology that are known before the judge is appointed.[263] This Section considers a range of predictive options as well as assesses the performance of prior presidents in appointing like-minded judges.[264] Further, the differences in the circuits of the U.S. courts of appeals are considered as factors that should be addressed in making predictions about judicial ideology.[265]

1. Appointing President and Ideology

Of the judges in the data sample, it might be contended that certain presidents were more or less successful in appointing judges to match their respective ideologies.[266] In particular, Presidents Ronald Reagan and George W. Bush have been targeted as having placed "ideologues" on the federal courts.[267] The data does show a clear statis-

---

[262] *See* Fischman & Law, *supra* note 7, at 198.

[263] *See id.* at 172.

[264] The Judge Database included a variety of demographic and biographical factors for each judge. Linear and logistic regressions were performed, as appropriate for the following characteristics of judges: rating by the American Bar Association during the nomination process; age at the time the judge was appointed; the political composition of the Senate; birth year; race; district court experience; state court experience; Justice Department experience; attorney general office experience; U.S. Attorney Office experience; Solicitor General office experience; law professor experience; private law firm experience; year appointed; Congress term during confirmation; whether the judge had taken senior status; and whether the judge was appointed during a recess appointment. In every case, the data did not support a statistically significant relationship between the Ideology Score and the listed factors. Because the regression analysis was limited to the 177 judges with sufficient sample sizes, it is possible that with more data for more judges, some of the factors might be correlated with Ideology Scores upon further study.

[265] *See* Fischman & Law, *supra* note 7, at 150.

[266] *See, e.g.*, Nadine Strossen, *The Current Assault on Constitutional Rights and Civil Liberties: Origins and Approaches*, 99 W. Va. L. Rev. 769, 808 n.182 (1997) (comparing the judicial appointments of Presidents Bill Clinton, Ronald Reagan, and George W. Bush); Scott Nance, *George W. Bush Court Appointments Emphasized Ideology over Diversity*, On the Hill, June 28, 2009, at A1 ("The judicial appointments of former president George W. Bush suggests [sic] that his motivation for appointing nontraditional judges was driven more by ideology and strategy than concerns for diversity . . . .").

[267] *See* Nance, *supra* note 266; Strossen, *supra* note 266, at 808 n.182.

tically significant correlation between the appointing president and Ideology Scores.[268] Yet the data did not support the hypothesis that Presidents Reagan and George W. Bush were uniquely responsible for appointing ideological judges. Figure 14 illustrates the average ideology of the judicial votes of the appointees of the last five presidents for the 177 judges with adequate samples.

Figure 14 - Ideology Scores by Appointing President (p=0.0003)
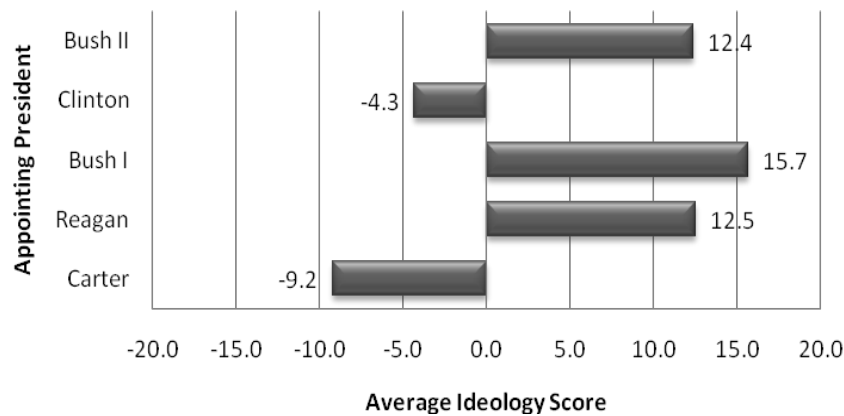
Of the appointees of the three Republican presidents who are currently serving on the U.S. courts of appeals, President George H.W. Bush's score stands out as being the highest. Notably, all three Republican presidents and President Carter appointed more ideological judges than President Clinton. Nevertheless, given the few presidents in the sample, it is difficult to infer any larger theory about the nature of the politics of presidents in regards to their appointed judges. Further, because many of the appointed judges for the older presidents have retired, the sample of appointees for those presidents is not random. As the time of appointment moves further back in time, the validity of the results is potentially lessened.

---

[268] The statistically significant correlation was a value of p=0.0003.

2. Circuit and Ideology

In studying or controlling for the ideology of judges, it is important to recognize that the different circuits have varied ideologies.[269] This could simply be due to the politics of the judges on the circuit or cultural factors among the judges such as a preference for consensus.[270] Regardless of the cause, it is essential to be able to identify whether results in future studies are the product of circuit ideology instead of the theorized variables.[271]

Among the regional circuits, there have emerged reputations for certain circuits being more liberal or conservative than others.[272] In particular, the Ninth Circuit is regularly labeled as liberal[273] while the Fourth,[274] Fifth,[275] and Eleventh[276] Circuits are considered to be conservative. The specific predictions based upon the conventional wisdom of the various circuits are only partially supported by the data as illustrated in Figure 15.

---

[269] *See* Fischman, *supra* note 162, at 3.

[270] *See id.*

[271] *See id.*

[272] *See, e.g.*, Bettina Boxall, *Bush-Era Road Rule on Forests Scrapped*, Chi. Trib., Aug. 6, 2009, at 34 (describing the Ninth Circuit as liberal); Bill Rankin, *Injured Deputy's Suit Thrown Out*, Atlanta J.-Const., Aug. 27, 2008, at 7C (describing the Fourth Circuit as conservative); Charlie Savage, *Obama Backers Fear Opportunities to Reshape Judiciary Are Slipping Away*, N.Y. Times, Nov. 15, 2009, at A20 [hereinafter Savage, *Obama Backers*] (describing the Eleventh Circuit as conservative); David G. Savage, *California Death Row Case Back in Court*, Chi. Trib., Nov. 2, 2009, at 13 [hereinafter Savage, *Death Row Case*] (describing the Fifth Circuit as conservative).

[273] *See* Boxall, *supra* note 272.

[274] *See* Savage, *Obama Backers, supra* note 272.

[275] *See* Savage, *Death Row Case, supra* note 272.

[276] *See* Rankin, *supra* note 272.

Figure 15 - Average Ideology Score by Circuit
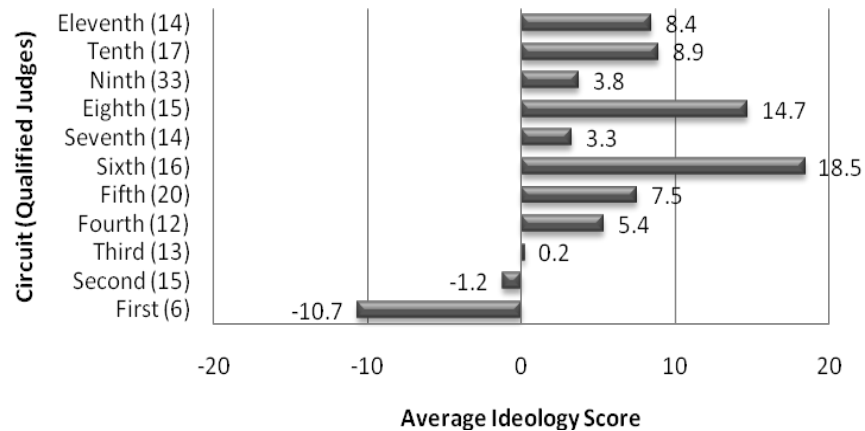
Although the Ninth is the fifth most liberal circuit, it does not appear to be uniquely liberal. Similarly, the Fourth, Fifth and, Eleventh Circuits have conservative scores, but they are not the most conservative. Indeed, the Sixth and Eighth Circuits are much more conservative on average than any of the three circuits with right-wing reputations. The result from the Fourth Circuit should not be terribly surprising as the reputation of the Fourth Circuit was in part based upon a time when there were notorious conservatives on the bench.[277] It is reasonable that the turnover in judges could account for a leftward move in the circuit's net ideology.[278] The First Circuit's score appears to be a significant liberal outlier. But this is almost surely due to the small sample size from that circuit. The First Circuit has by far the smallest docket of any of the circuits and only had 278 cases in the data set. Additionally, because of the limited number of qualifying judges (six), the interactions between judges were not as varied as the other circuits. Further research is warranted to determine if the First Circuit judges are as liberal as the Ideology Scores indicate.

Another significant consideration in researching judicial ideology at the federal appellate level concerns the homogeneity of circuits in

---

[277] *See* Larry Margasak, *Court Nominee Survives Senate Test*, Bos. GLOBE (Nov. 18, 2009), http://www.boston.com/news/nation/washington/articles/2009/11/18/court_nominee_survives_senate_test ("Last week, the Senate confirmed US District Judge Andre Davis . . . giving Democratic nominees a 6-to-5 edge on the Fourth Circuit that once was a conservative legal bastion.").

[278] *See id.*

terms of ideology. Although the net ideology numbers paint part of the picture, the distribution of the judges along the ideological spectrum is also significant. In Figure 16, the average score differential (in either ideological direction) for each circuit for the judges in the 177-judge sample is indicated.

Figure 16 - Average Ideological Variance by Circuit
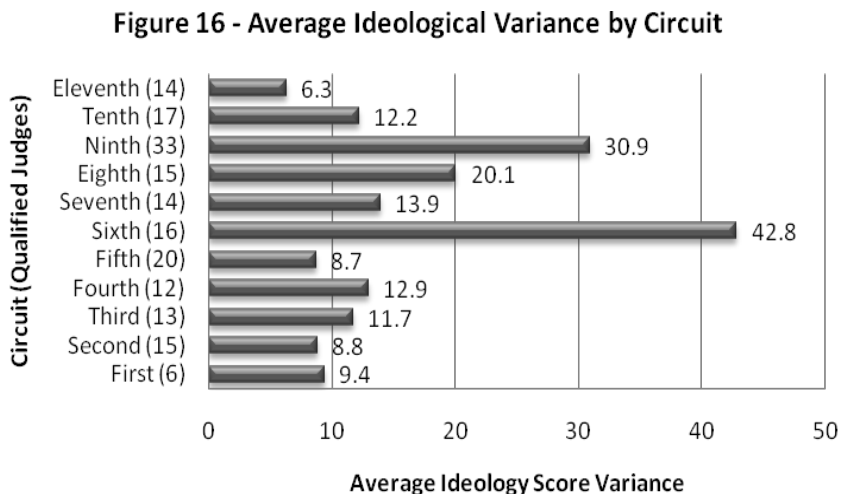
Based upon the data, the Sixth, Eighth, and Ninth Circuits appear to have sharp ideological divisions between judges. In contrast, the Second, Fifth, and Eleventh Circuit judges are largely ideologically homogenous. The variance cannot be explained merely by circuit size. Indeed, the Fifth Circuit and Ninth Circuit have the two largest sets of judges in the 177-judge sample and are on opposite ends of the variance spectrum. The finding that the Ninth and Sixth Circuits are especially polarized by ideology is consistent with the anecdotal indications that those circuits have been plagued by ideological divides among the judges.[279]

3. Law School and Ideology

Reviewing biographical factors related to judicial ideology might uncover certain factors that are related to future judicial ideology. A common refrain of conservative critics of the federal judiciary is to

---

[279] *See* Boxall, *supra* note 272 (describing the Ninth Circuit as having "a liberal reputation"); R. Jeffrey Smith, *Ideological Warfare Rages on Federal Appeals Courts*, WASH. POST, Dec. 8, 2008, at A1 ("Ideological trench warfare is frequently on display in the 6th Circuit's austere fourth-floor hearing room . . . which shifted to Republican-appointee control in mid-2005.").

blame "liberal judicial activism" on Ivy League educations.[280] Judges on the courts of appeals actually come from a wide range of law schools making a test of such a hypothesis possible. A linear regression indicated that there is substantial evidence in the data to support such a claim (p=0.0151). Figure 17 illustrates how law school ranking as determined by the 2009 U.S. News & World Report Law School Rankings[281] are related to the Ideology Scores of the 177 studied judges.

Figure 17 - Ideology Score by Law School Attended (p=0.0151)
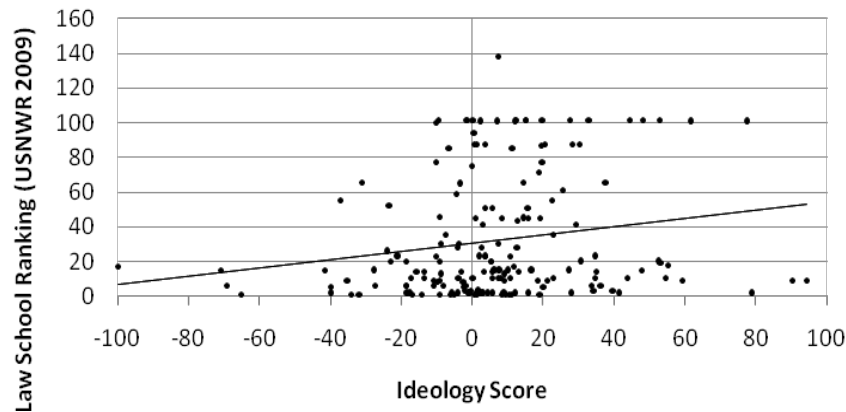
The correlation is unmistakable, as the data indicates that for approximately every ten "lower" ranks of law schools, a judge's Ideology Score increased (in a conservative direction) by forty points. Notably, a regression showed no support for the idea that private schools were more likely to be associated with liberal judges (p=0.0587). Further, the data indicated the same relationship in both appointees of Republican and Democratic presidents, as indicated in Figures 18 and 19.

---

[280] *See, e.g.*, Thomas Sowell, *The Great Danger of Supreme Quotas*, Human Events Online (July 7, 2005), http://www.humanevents.com/article.php?id=8006 (contending that "liberal judicial activism" is a product of an elite law school mentality).

[281] The *U.S. News & World Report* rankings are available at http://grad-schools.usnews.rankingsandreviews.com/best-graduate-schools/top-law-schools. Although the flaws in the *U.S. News & World Report* Rankings are many, it is simply the only ranking system to cover all of the schools in the data set. *See* Gary Blankenship, *Law Schools: Time to Get Practical?*, Fla. Bar News, Aug. 1, 2009, at A1 (describing the "disproportionate influence" of the rankings on legal education).

Figure 18 - Ideology Score by Law School Attended for Democratic Appointees
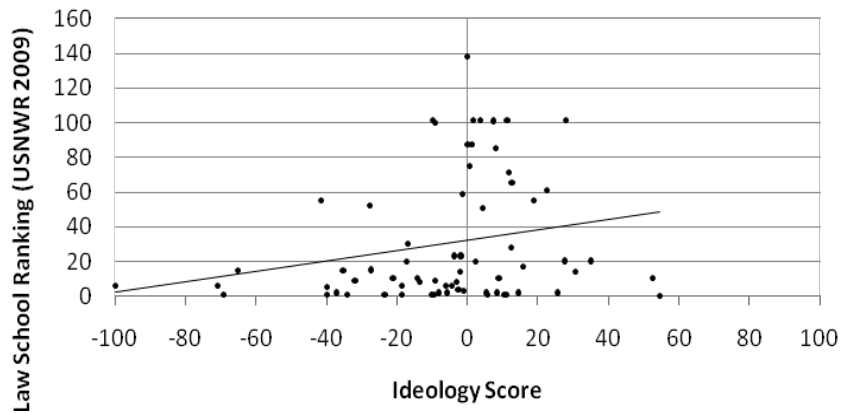
Figure 19 - Ideology Score by Law School Attended for Republican Appointees
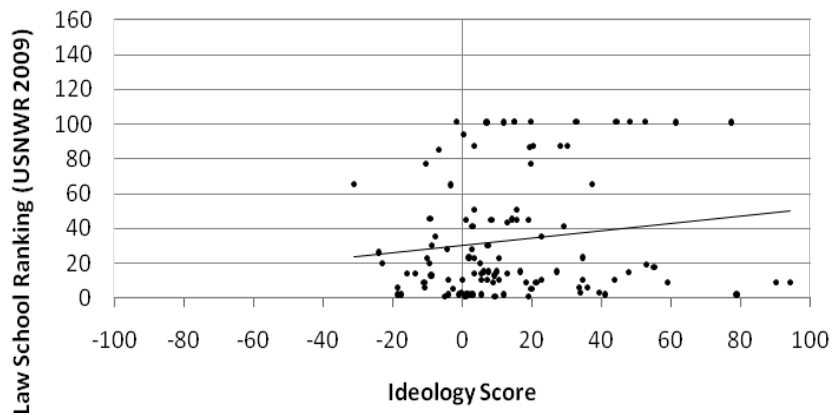
One possible explanation for the strong correlations indicated above is that the circuit of the judge is a variable responsible for both the ideology and law school attended. Theoretically, because elite law schools are concentrated in a few circuits, it is likely that certain circuits would have a higher concentration of elite law school graduates. If those circuits were also politically liberal, then the correlation might be explained entirely by the circuit and not the law school attended. A separate regression between the Ideology Scores and law school ranking

was performed while controlling for the judges' circuits, and the correlation was still statistically significant (p=0.0230). Although the data does not indicate a clear causation, the finding is significant enough to warrant further investigation.

4. Government Experience and Ideology

Although a variety of background factors were examined in relation to the Ideology Scores, government experience was the only other factor that showed a statistically significant relationship. It was hypothesized that because conservatism is more strongly associated with an aversion to government, judges who had been affiliated with government bodies (outside of the judiciary) would be more liberal than judges without such experience. A regression analysis on the data offered strong support for the hypothesis (p=0.0008). Figure 20 indicates that the effect was pronounced for both appointees of Republican and Democratic presidents.

Figure 20 - Ideology by Government Experience (p=0.0008)
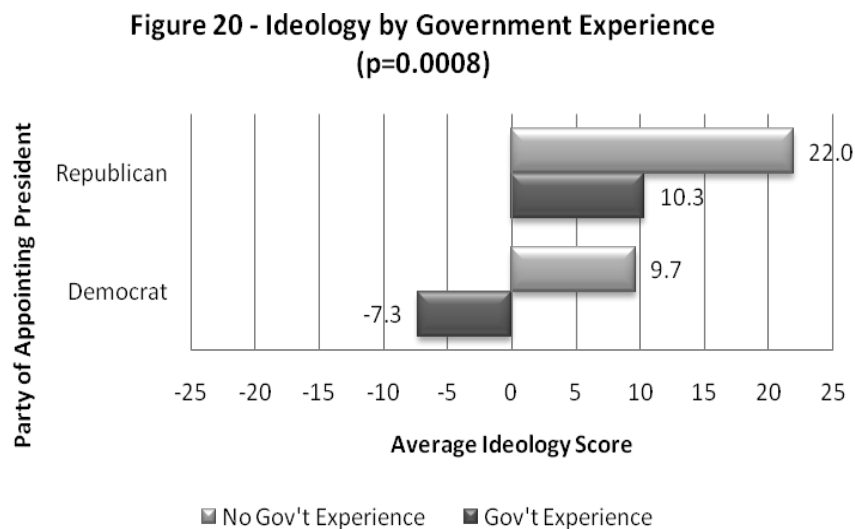
The correlation was dramatic enough to provide a very strong indicator of liberalism on the federal judiciary. Combined with law school attended, government experience provides a substantial hint about the characteristics that might be strongly connected to the politics of a judge on the U.S. courts of appeals. Certainly, a Democratic president might be worried about a graduate of a lower-ranked law school with no government work experience. Similarly, the data indicate that a Re-

publican president should be wary of elite law school graduates who have worked in the public sector.

IV. Data Analysis

The results above indicate some new findings about the nature of ideology in the U.S. courts of appeals. Nevertheless, there are reasons to be cautious about broad conclusions. As with any empirical study, it is helpful to examine reliability, validity, and potential limitations of the data. Each of those areas of concern is discussed below.

A. *Reliability*

Reliability is the degree to which the measurement would yield the same results when applied by others.[282] Because this study is the first to systematically analyze judicial ideology among federal judges at the federal appellate level based upon actual performance on the bench with a unique data set, reliability cannot be determined by comparison to other empirical studies.[283] Instead, reliability is evaluated by the quality of the coding and analysis.[284]

Case data was acquired from LexisNexis for each of the circuits studied. Some of the objective data (such as party names, citation, and opinion date) was harvested directly from downloads of cases using a proprietary computer software program.[285] The program has been used previously in other studies reviewing federal appellate court opinions.[286] The remaining data was coded by law students and law library staff. Four of the students were solely responsible for the Case Database and one created the biographical information database.

After an initial coding of the key variables in the Case Database, a second coding of those variables was performed for random samples of the data by persons other than those who did the original coding for the individual cases, yielding acceptable inter-coder reliability levels for

---

[282] Epstein & King, *supra* note 124, at 83 ("Reliability is the extent to which it is possible to replicate a measurement, reproducing the same value (regardless of whether it is the right one) on the same standard for the same subject at the same time.").

[283] *See id.*

[284] *See id.*

[285] *See* David L. Schwartz, *Courting Specialization: An Empirical Study of Claim Construction Comparing Patent Litigation Before Federal District Courts and the International Trade Commission*, 50 Wm. & Mary L. Rev. 1699, 1735 n.174 (2009) (searching for ITC cases using Westlaw and LexisNexis); Schwartz, *Practice Makes Perfect?*, *supra* note 137, at 239 (locating Rule 36 cases through a Westlaw search).

[286] *See* Schwartz, *supra* note 285, at 1735 n.174; Schwartz, *supra* note 137, at 239.

the tested variables. In addition, a variety of checks were performed to ensure internal consistency of variables that were necessarily interconnected.[287]

B. *Validity*

Validity is the degree to which the measurement used in an empirical study reflects the concept measured.[288] The assessment of the study is more complex for validity than for reliability.[289] Generally, validity can be understood along a variety of axes.[290] For example, Professors Lee Epstein and Gary King, in their call for improved empirical legal scholarship, identified three possible ways to view validity: "facial validity, unbiasedness, and efficiency."[291] Although establishing each of those categories is unnecessary, it is helpful to keep them in mind when reviewing a study trying to measure a complex concept like "judicial ideology."[292] Each of those categories is discussed below, but, initially, it is important to first return to the discussion of what exactly this study is measuring.

1. Concept Measured

As detailed in Part I.A, "judicial ideology" can be a slippery concept that is tricky to capture in empirical studies.[293] Consequently, it is helpful to revisit the topic of whether the results above are truly indicative of judicial ideology. There are at least two reasons to believe that Ideology Scores describe judicial ideology.

First, the success of the scores in predicting case outcomes against the current dominant models is strong evidence that the study is properly directed. Even without the relative comparisons with other models, the Ideology Scores were shown, based upon the data, to have statistically significant relationships with outcomes in civil rights cases.

---

[287] For example, the party labels in the coding includes "criminal defendant." In order for there to be a criminal defendant, the outcome for the prevailing party cannot be "civil plaintiff" or "civil defendant." A variety of cross-checks were employed to quality check and correct errors within the data set.

[288] Epstein & King, *supra* note 124, at 87 ("Validity is the extent to which a reliable measure reflects the underlying concept being measured.").

[289] *See id.*

[290] *See id.*

[291] *Id.* at 89.

[292] *See id.* ("[N]o one of these is always necessary, and together they are not always sufficient, even though together they are often helpful in understanding when a measure is more or less valid.").

[293] *See* Fischman & Law, *supra* note 7, at 156.

Second, although this Article does not deploy the same methodology as the Martin-Quinn scores, it relies on the same basic assumptions about determining ideology: agreement and disagreement between judges is indicative of shared values.[294] With strong prior assessments of some judges, those values can be mapped onto an ideological scale.[295] Insofar as the Martin-Quinn scores have become the leading measure of U.S. Supreme Court justice ideology, the validity of the Ideology Scores should be similarly well-grounded.[296]

2. Facial Validity

As Professors Epstein and King noted, "[a] measure is facially valid if it is consistent with prior evidence, including all quantitative, qualitative, and even informal impressionistic evidence."[297] Even though this Article contends that the model outlined herein offers a viable alternative to prior measures of judicial ideology, for validity purposes, it is helpful to consider how those values relate to each other.[298] Linear regressions were performed to determine if there was a statistically significant relationship between the Ideology Scores and both the Common Space Scores and ideology of the appointing president. In both instances, such a relationship was supported by the data.[299]

Further, as illustrated in Appendix A, a great many of the judges fit expectations of ideology based upon their appointing presidents. Still, it is helpful to consider some of the higher profile judges in regards to their Ideology Scores. Figure 21 below indicates, for thirteen notable judges, the Ideology Scores and scaled Common Space Scores.

---

[294] *See* Jacobi & Sag, *supra* note 21, at 3–4.

[295] *See* Fischman & Law, *supra* note 7, at 162–63.

[296] *See* Jacobi & Sag, *supra* note 21, at 3.

[297] *Id.*

[298] *See* Epstein & King, *supra* note 124, at 87.

[299] p=0.000 in both cases.

Figure 21 - Ideology and Common Space Scores for Notable Judges

Legend: Common Space Score (Scaled) — Ideology Score
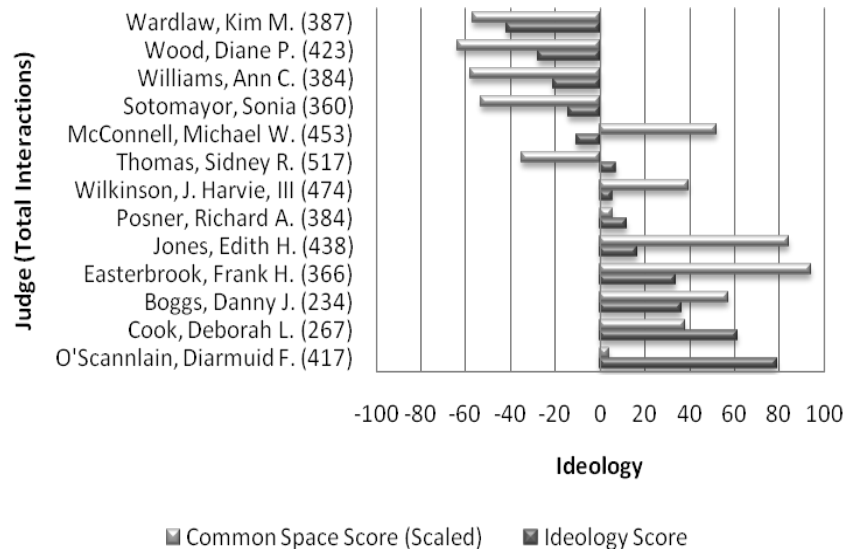
Among the thirteen judges listed in Figure 21, there are six judges who were rumored to be Supreme Court candidates for Presidents Bush and Obama (Judges Wardlaw, Wood, Cook, Williams, Thomas, and Jones),[300] one confirmed justice (Justice Sotomayor), one well-known social conservative who recently left legal academia (Michael McConnell), and five prominent conservative jurists (Judges Posner, Wilkinson, Boggs, O'Scannlain, and Easterbrook).[301] In most instances, their Ideology Scores fall on the side of the spectrum with which they are strongly associated. As might be expected, Judges Easterbrook and O'Scannlain, who are conservative icons on their respective circuits, are amongst the

---

[300] Stuart Taylor Jr. & Evan Thomas, *Keeping It Real*, Newsweek, Nov. 14, 2005, at 22 (contrasting Judge Jones, as a potential Supreme Court appointee, to Justice Alito); Bash, *supra* note 2.

[301] Peter Slevin, *Was Blogger Threatening Judges, or Just Exercising Free Speech?*, St. Paul. Pioneer Press, Aug. 16, 2009, at A4 (reporting threats made against Judges Easterbrook, Posner, and Bauer on the Seventh Circuit); Carol J. Williams, *Three Key Players Reflect the 9th Circuit's Ideological Span*, L.A. Times, Apr. 19, 2009, at A13 (describing the three most prominent judges on the Ninth Circuit as Judges Kozinski, Reinhardt, and O'Scannlain); Benjamin Wittes, *The Best Judges He Can't Pick*, Wash. Post, May 3, 2009, at B01 (discussing potential replacements for Justice Souter on the Supreme Court).

most conservative judges in the data set.[302] Of significance, Judge O'Scannlain appears as conservative using the Ideology Scores, in accordance with his reputation, whereas the Common Space Scores considered him to be the most moderate of all of the Republican appointees in the data set.

Another interesting example, not included above, is Judge Richard Tallman of the Ninth Circuit. President Clinton nominated Judge Tallman even though it was known that Judge Tallman was a Republican.[303] Because the Common Space Scores assume every Clinton nominee was liberal, his Common Space Score was –0.327. Using the Ideology Score method, however, Judge Tallman was the most conservative judge appointed by a Democratic president, having a score of 54.7. At least in the case of Judge Tallman, the information about his background politics was borne out in the Ideology Scores.

Notably, President Obama's decision to nominate Sonia Sotomayor to the Supreme Court may have resulted in the selection of a moderate liberal among frequently-mentioned candidates.[304] Although Judges Wardlaw, Wood, and Williams were further from the midpoint than Justice Sotomayor, Justice Sotomayor was a clear liberal in 2008 according to the Ideology Scores. At this point in time, however, there is no data connecting a judge's ideology based upon appellate court behavior with his or her ideology on the Supreme Court. Indeed, if one subscribes to some version of the strategic model of judging, an appellate judge may intentionally moderate his or her views on the appellate court to increase the chances of nomination to the Supreme Court.[305]

3. Unbiasedness

According to Epstein and King's categorization of validity tests, "[a] measurement procedure is unbiased if it produces measures that are right on average across repeated applications."[306] Because this study re-

---

[302] *See* Slevin, *supra* note 301; Williams *supra* note 301.

[303] Henry Weinstein, *Court to Consider Delay of Recall Vote*, L.A. Times, Sept. 20, 2003, at A1.

[304] *See* Charlie Savage, *Conservatives Map Strategies on Court Fight*, N.Y. Times, May 17, 2009, at A1 (discussing President Obama's contemplated replacements for Justice Souter).

[305] *See* Miriam A. Cherry & Robert L. Rogers, *Tiresias and the Justices: Using Information Markets to Predict Supreme Court Decisions*, 100 Nw. U. L. Rev. 1141, 1155 (2006) ("The strategic model . . . suggests that judges act in ways that maximize their incentives and their prestige."); Richard A. Posner, *What Do Judges and Justices Maximize? (The Same Thing Everybody Else Does)*, 3 Sup. Ct. Econ. Rev. 1, 5 (1993) ("Supreme Court Justices are often appointed from the ranks of federal court of appeals judges and although the probability of such an appointment is low . . . it figures in the thinking of some judges.").

[306] Epstein & King, *supra* note 124, at 92 (emphasis omitted).

lies on "revealed preferences" of the voting blocs combined with a formal rule that is rarely in dispute (standard of review), the risk of bias in measurement should be less than relying on other techniques.[307] The standard of review is often briefed by both parties, and in a random sample of the briefs in one hundred cases in the Case Database, there were no instances where the parties had a dispute as to the applicable standard(s) of review. Further, the Ideology Scores appear to avoid some of the problems of biased coding of case outcomes in regards to political preferences of the judges.[308]

4. Efficiency

The last Epstein and King test for validity is that "[e]fficiency helps us choose among several unbiased measures, with the basic idea being to choose the one with the minimum variance."[309] In relation to the Common Space Scores and the party of the appointing president, there is reason to believe that the Ideology Scores are more efficient.[310] Typically, if a particular measure omits relevant information, it is less efficient because there will be less consistency in application.[311] Each of the three measures omits some relevant information.[312] Ideology Scores do not factor in the political factors at the time of nomination. In contrast, Common Space Scores and the party of the appointing president do not include any information regarding a judge's behavior on the bench.[313] Rather than omitting such information, it seems reasonable to believe that inclusion of significant performance data, as opposed to political

---

[307] *See id.* at 94 ("So, instead of (or sometimes in addition to) asking respondents to answer research questions directly, it is usually better to look for revealed preferences, which are consequences of theories of motive that are directly observable in real behavior.") (emphasis omitted).

[308] *See* Edwards & Livermore, *supra* note 24, at 1924–25 ("Some studies seek to code case outcome according to topical or political criteria. For example, in an empirical study done by Glendon Schubert, Supreme Court decisions issued between 1946 and 1963 were coded along two axes—political liberal/conservative and economic liberal/conservative. Cases also have been coded as pro-/antienvironment, pro-/anticriminal defendant, pro-/anti-civil rights, and so on. Perhaps the most common metric used in empirical studies is a simple 'left/right' or 'liberal/conservative' binary. These topical or political measures used to describe cases will necessarily simplify a court's holding and reduce what may be a complex and nuanced decision into an often uninformative binary.").

[309] Epstein & King, *supra* note 124, at 95.

[310] *See id.* at 95–96.

[311] *See id.* at 96 ("The result is that any one application of the measure with more information will be likely to yield an answer closer to the truth than any one application of the measure with less information.").

[312] *See id.* at 95–96.

[313] *See id.*

factors during the nomination process, in the Ideology Scores would produce a more efficient model of judicial ideology.[314] This conclusion is borne out by the predictive success of the Ideology Scores versus the other two measures.

C. *Limitations of the Data*

There are several limitations to the data utilized in this study. Consequently, as with any empirical study, it is important to recognize those limitations so that inferences from the data are properly drawn. Because this is the first study using this newly created data set, the need to carefully analyze the confines of the data is even more important. There are several ways in which the data used herein is notably limited.

1. Time Limitations

All the data studied in this Article were from opinions issued in 2008. This has several implications. First, it is uncertain whether the judges studied have always had the same ideology or if their ideology has changed over time. Similarly, going forward, the judges may drift in their ideologies.[315] Second, the results should not be used to create inferences about all of the judges appointed by a particular president. Especially for presidents that served decades ago, the remaining appointees who were still issuing opinions in 2008 may not be representative of the overall population of judges appointed by those presidents.

2. Data-Gathering Limitations

The study also excluded opinions that did not use language relevant to a standard of review because of the way the LexisNexis searches were executed.[316] That omission means that a substantial portion of opinions by judges were not considered in this study. If the sample used herein were a random sample of the overall opinion population, the omission would not be a statistical problem. It is possible, however, that in the excluded cases, had the judges included a standard of review, there might have been different results that could have altered conclusions derived from the data of this study. Nonetheless, if the levels of excluded opinions were relatively consistent among judges, the order of the Ideology Scores should not substantially change.

---

[314] *See id.*

[315] *See* Epstein et al., *supra* note 145, at 1540.

[316] *See* Motomura, *supra* note 186, at 474; Scott, *supra* note 187, at 305.

Another limitation of the data concerns the degree to which certain mixes of case types might distort the results. The various regressions that were run in this study controlled for the case mix between criminal and civil cases.[317] Nevertheless, it is possible that finer distinctions in composition of caseload per circuit and per judge could account for some of the variation in Ideology Scores. Without more data, it is impossible to assess the significance of this limitation.

3. Coding Limitations

As this study relied entirely on a metric that measured certain aspects of the opinion itself, some potential problems emerge. An initial problem with such metrics is that they rely on judicial rhetoric that might only be window dressing for the real basis for the opinion.[318] As Judge Posner has commented, examining the rhetoric of judicial opinions to determine what a judge is really doing will often mislead the observer.[319] Because the measure used herein takes a judge's words at face value, this is an applicable concern.[320] Standards of review, however, do not fit the normal rubric of potentially misleading rhetoric.[321] The standards are based upon long lines of cases, are usually agreed upon in the competing party briefs, and do not determine the case outcome alone.[322] As a result, this study does not raise significant concerns about overreliance on judicial rhetoric.[323]

A variation on this objection stems from the inherent reduction of an opinion to quantifiable data that occurs in a model like the one used herein.[324] Because judicial opinions are not written with a particular formula, the difficulty of coding different styles and forms should not be underestimated.[325] Reducing lengthy opinions to a couple of variables may miss a substantial amount of nuance in the opinion.[326] This criticism is prominent among judges who have been disconcerted with using quantitative analysis to understand judicial decision mak-

---

[317] *See supra* notes 192–199 and accompanying text.

[318] *See* Richard A. Posner, *The Jurisprudence of Skepticism*, 86 Mich. L. Rev. 827, 865 (1988) ("We should not be so naive as to infer the nature of the judicial process from the rhetoric of judicial opinions.").

[319] *See id.*

[320] *See id.*

[321] *See* Cross, *supra* note 72, at 1500.

[322] *See id.*

[323] *See id.*

[324] *See* Fischman & Law, *supra* note 7, at 156.

[325] *See id.*

[326] *See id.*

ing.[327] In this case, however, the key variables were relatively easy to code and objective in nature. Disposition and standards of review are part of the formula for many opinions, and the results rely almost exclusively on those variables. Spotting those portions of the opinion does not rely on any formulaic conception of the opinion itself.

4. Selection Effects

As with any project that studies court cases that are included in a sample based upon the actions of third parties (i.e., litigants), there is a concern about selection effects.[328] A selection effect is "a causal relationship between the distribution of disputes and other variables of litigation."[329] There are many points during litigation when a selection effect could occur, including pre-filing, pretrial, during trial, pre-verdict, post-verdict, pre-appeal, during appeal, and post-appeal.[330] If a selection effect distorted the case mix in a way important to the study, it would call into question the study's validity.[331]

Because this study is exclusively focused on the behavior of judges on the U.S. courts of appeals, the need to account for certain selection effects is limited.[332] Selection effects are only significant for this study if they distort the case mixes of individual judges or circuits relative to other judges or circuits.[333] Otherwise, as long as the selection effects are *consistent* among the units of measure, they should not implicate the validity of the Ideology Scores.

---

[327] *See* Edwards & Livermore, *supra* note 24, at 1910.

[328] *See* Keith N. Hylton, *Asymmetric Information and the Selection of Disputes for Litigation*, 22 J. LEGAL STUD. 187, 203–05 (1993) ("[T]he disputes that are most likely to influence legal doctrine tomorrow are those in which the divergence in litigants' expectations . . . is greatest . . . . This imbalance should influence the menu of issues reaching appellate courts."); George Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1, 2 (1984) ("It is well known, however, that only a very small fraction of disputes comes to trial and even a smaller fraction is appealed . . . . It is very difficult to infer specific characteristics from observations of [appeals cases], especially where there is no evidence that the observations (the disputes selected for appeal) were selected randomly."); Cass Sunstein, *Judging National Security Post-9/11: An Empirical Investigation*, 2008 SUP. CT. REV. 269, 271 ("A central point here involves selection effects: Because litigants are responsive to the likelihood of victory, it is important to be careful in drawing lessons from any particular invalidation rate.") (emphasis omitted).

[329] Kate Stith, *The Risk of Legal Error in Criminal Cases: Some Consequences of the Asymmetry in the Right to Appeal*, 57 U. CHI. L. REV. 1, 19 n.55 (1990).

[330] *See id.* at 19–28.

[331] *See id.*

[332] *See id.* at 25–28.

[333] *See id.*

At the most basic level, selection effects rely on theoretical conceptions of the judiciary about the incentives of litigants.[334] Under the Priest-Klein hypothesis, one would expect appellate outcomes to split 50% between affirmances and reversals since the parties would settle appropriately to avoid other outcomes.[335] In the data set in this study, however, and consistent with prior examinations of the Priest-Klein hypothesis, the results do not support the hypothesis in its broadest form because affirmances occur at a much higher rate than 50% regardless of the standard of review used. Importantly, in criminal cases, the settlement (plea) structure creates different incentives such that a 50% split is unlikely to occur. Nevertheless, even in the non-criminal cases within the sample, the affirmance rate was 72.2%, far higher than the expected equilibrium rate.

The explanation for the difference between the observed rates and the Priest-Klein hypothesis also addresses the issues raised by selection effects more broadly. In the appellate environment, in many instances, the marginal cost of appeal is low compared to that of a trial.[336] Further, there is significant uncertainty in predicting appeal outcomes.[337] Beyond the issues being legally close,[338] the parties will not know who composes the panels until shortly before the oral argument.[339] By that time, the briefing was long since completed, and any settlement is unrealistic.[340]

---

[334] *See* Sunstein, *supra* note 328, at 271.

[335] *See* Priest & Klein, *supra* note 328, at 5. A variety of studies have looked at whether the Priest & Klein hypothesis is supported by empirical evidence. *See, e.g.*, Theodore Eisenberg, *Testing the Selection Effect: A New Theoretical Framework with Empirical Tests*, 19 J. Legal Stud. 337, 342–47 (1990); Randall S. Thomas & Kenneth J. Martin, *Litigating Challenges to Executive Pay: An Exercise in Futility?*, 79 Wash. U. L.Q. 569, 590 (2001); Robert E. Thomas, *The Trial Selection Hypothesis Without the 50 Percent Rule: Some Experimental Evidence*, 24 J. Legal Stud. 209, 226–27 (1995).

[336] Meehan Rasch, *Not Taking Frivolity Lightly: Circuit Variance in Determining Frivolous Appeals Under Federal Rule of Appellate Procedure 38*, 62 Ark. L. Rev. 249, 264 (2009) ("[E]motion-laden, nonprevailing parties have little to lose by appealing, especially given the minimal court costs associated with taking an appeal.").

[337] *See* Richard L. Revesz, *Litigation and Settlement in the Federal Appellate Courts: Impact of Panel Selection Procedures on Ideologically Divided Courts*, 29 J. Legal Stud. 685, 688–89 (2000).

[338] *See* Brian Z. Tamanaha, *The Distorting Slant in Quantitative Studies of Judging*, 50 B.C. L. Rev. 685, 748 (2009) (quoting Cass R. Sunstein et al., Are Judges Political?: An Empirical Analysis of the Federal Judiciary 16 n.20 (2006)) ("As the authors acknowledge, the subset of cases that are actually appealed following trial are more likely to have 'a degree of indeterminacy in the law.'").

[339] *See* Revesz, *supra* note 337, at 688–89 ("With one exception, the United States Courts of Appeals announce the composition of their panels only shortly before the oral argument, typically after all the briefs have been filed. Panels are announced 1 week before the argument in the First Circuit, on the Thursday before the argument in the Second Circuit, 10

There is no effective way to be sure that there are not exogenous variables or selection bias issues that have distorted the results.[341] This problem is not unique to this study, as this is a concern in virtually every study of federal judges.[342] Nonetheless, because the Ideology Scores outperformed the existing measures as described in Part III, whatever effects that emerged from selection issues did not fully diminish the overall value gained by using the measure described herein.

5. Unidimensionality

There is inevitably a concern when modeling judicial ideology that one dimension along a liberal-to-conservative continuum might be insufficient to capture the nuances of the concept.[343] The unidimensionality problem afflicts case outcome coding, external proxy measures, and agnostic measures.[344] Although it is theoretically possible to use, for example, agnostic coding along multiple dimensions, the lack of agreement about the contents of those dimensions has inhibited such approaches.[345] Consequently, the unidimensionality problem is a concern regardless of which of the three methodologies is used.[346]

The unidimensionality problem could appear in a variety of ways.[347] For example, it is possible to conceive of a judge who is libertarian, distrusting government in all cases and favoring individual liberty.[348] This hypothetical judge might appear liberal in criminal cases by voting against the government while appearing conservative in civil cases because he or she has classically liberal economic views.[349] A unidimensional spectrum fails to treat such a judge any differently than one who

---

days before the argument in the Third Circuit, on the day of the argument in the Fourth Circuit, 1 week before the argument in the Fifth Circuit, 2 weeks before the argument in the Sixth Circuit, on the day of the argument in the Seventh Circuit, approximately 1 month before the argument in the Eighth Circuit, on the Monday of the week before the argument in the Ninth Circuit, 1 week before the argument in the Tenth Circuit, 1 week before the argument in the Eleventh Circuit, and on the day of the argument in the Federal Circuit.").

[340] *See id.*

[341] *See* Priest & Klein, *supra* note 328, at 2.

[342] *See id.*

[343] *See* Fischman & Law, *supra* note 7, at 150–54.

[344] *See id.* at 164 ("But the agnostic approach to coding has its shortcomings as well. For starters, models of this type generally assume a one-dimensional spectrum, yet it may be difficult to know what distortions this assumption might cause.").

[345] *See* KEITH T. POOLE, SPATIAL MODELS OF PARLIAMENTARY VOTING 141–47 (2005).

[346] *See id.*

[347] *See* Fischman & Law, *supra* note 7, at 150–54.

[348] *See id.*

[349] *See id.*

is pro-government in criminal cases, but favors government regulators in some civil matters. The assumption of unidimensionality can be particularly troublesome when studies examine cases over a large period of time.[350] What constitutes "liberal" or "conservative" in 1850, 1900, 1950, and 2000 might vary substantially.[351] So, even the one dimension identified by researchers as the basis for their study might not be static over time

Yet, there is evidence from prior studies that, in fact, a unidimensional model accounts for a great deal of judicial behavior in American courts.[352] At the U.S. Supreme Court level, a study by Professors Bernard Grofman and Timothy Brazill found that a single dimension would predict an estimated eighty to ninety-three percent of voting by the Supreme Court from 1953 to 1991.[353] In particular, the study found the unidimensional model worked especially well with more recent Courts.[354] This should not be terribly surprising because the American political system is based upon two parties.[355] Those parties aggregate a collection of policy views in contrast to the other party.[356] Because the President appoints the judges and Senate confirms them, it follows that the appointed judges embody, to a large degree, the political ideology of the political figures that supported their nomination. It is unclear, however, if judges on lower courts exhibit the same degree of belief cohesiveness to fully disregard the unidimensionality problem.[357] Nonetheless, the observed high degree of correlation among various beliefs underlying ideology makes this worry less significant than it initially

---

[350] *See id.*

[351] *See id.*

[352] *See* Jacobi & Sag, *supra* note 21, at 17 ("[T]here is considerable evidence that a single dimension captures the vast majority of judicial behavior. The fact that one dimension captures most judicial behavior does not imply that legal analysis is simplistic, only that most judicial considerations are generally quite highly correlated. For instance, a Justice's view on breadth of congressional power is likely to be closely correlated to that Justice's view on states rights and on less directly related issues, such as abortion or the death penalty. Thus we can speak of a Justice being conservative or liberal, and although occasionally the justices surprise their audience, the majority of a Justice's decisions are consistent with these expectations. In fact, Justices' votes have been shown to be highly predictable across a number of issue areas including the death penalty, freedom of speech, search and seizure, federalism, and administrative law.").

[353] *See* Bernard Grofman & Timothy J. Brazill, *Identifying the Median Justice on the Supreme Court Through Multidimensional Scaling: Analysis of "Natural Courts" 1953–1991*, 112 Pub. Choice 55, 58 (2002).

[354] *See id.*

[355] *See* Giles, *supra* note 11, at 627.

[356] *See id.*

[357] *See* Jacobi & Sag, *supra* note 21, at 19–20.

might appear.[358] Notably, the performance tests using civil rights cases, as described earlier, provide further support for using a unidimensional scale.[359] Even when excluding every civil rights case in the data set, Ideology Scores based upon the non-civil-rights cases were able to better predict the outcomes of civil rights cases as opposed to competing measures.

Conclusion

This study expands the metrics of judicial ideology for courts below the U.S. Supreme Court to techniques based upon the behavior of judges. The Ideology Scores and the techniques underlying their computation appear to be a significant advance over existing measures of the concept. In applying these scores, the results of the study were as follows. The eleven studied circuits exhibited very different average ideologies and heterogeneity of ideology. The data did not support the hypothesis that Presidents Ronald Reagan and George W. Bush appointed uniquely ideological judges. Of the 177 judges studied in 2008, the data indicated that judges appointed by Republican presidents were more ideological than those appointed by Democratic presidents. Notably, prior government work experience and elite law school attendance were strongly correlated with political liberalism on the bench.

There are several avenues of new research that are suggested by the results of this Article. First, new data sets are desperately needed for federal courts below the Supreme Court. Although the Songer database is suitable for many projects, it is simply impossible to develop comprehensive measures of individual judges without greater sample sizes. New data also offers more opportunities to test various measures of judicial behavior without circularity problems.

Second, the techniques outlined in this Article can function just as well in assessing the ideology of district court judges. This study did not include scores for those judges primarily because of sample size concerns. Nevertheless, because district courts are often the gatekeepers of the federal court system, it would valuable to know what their judicial ideologies are.

---

[358] *See* Fischman & Law, *supra* note 7, at 151 ("In practice, however, the challenge that multidimensionality poses to the measurement of judicial ideology may not be as severe as this hypothetical example might suggest. The views that people hold across a range of questions tend to correlate with one another in systematic ways.").

[359] *See supra* notes 254–259 and accompanying text.

Third, new multi-level, multi-court research should be performed to allow more and different points of comparison for assessing judicial behavior. By continually focusing on single courts and court levels, empirical research may be overlooking crucial information about court interrelations. Indeed, as the research on panel composition effects demonstrated, the assumptions about the isolation of judicial behavior that existed for decades did not hold up when interactions were actually studied.

Ultimately, despite the numerous insights developed by empirical legal studies of the judiciary over the last few decades, there is substantial room for the field to grow. This Article hopefully contributes to the direction of that growth by harnessing information from the actual judicial opinions, interactions between judicial levels, and case dispositions into a single measure. Further, by focusing on courts below the Supreme Court, a greater picture of judicial behavior and new insights into members of America's life-tenured judiciary are gained.

Appendix A: Judicial Ideology Scores from the most conservative to most liberal for judges with at least 300 interactions with other judges.

Rank	Judge (Total Interactions)	Circuit	Common Space Score	Ideology Score
1	Rogers, John M. (366)	6	0.3920	96.4
2	McKeague, David W. (386)	6	0.5310	91.3
3	Batchelder, Alice M. (308)	6	0.5020	80.4
4	O'Scannlain, Diarmuid F. (417)	9	0.0230	72.6
5	Griffin, Richard A. (421)	6	0.5310	61.2
6	Gruender, Raymond W. (602)	8	0.3010	56.9
7	O'Brien, Terrence L. (345)	10	0.4405	56.1
8	Riley, William J. (573)	8	0.2880	49.7
9	Ikuta, Sandra S. (329)	9	0.5310	49.0
10	Canby, William C., Jr. (320)	9	-0.1650	46.7
11	Siler, Eugene E., Jr. (325)	6	0.3390	43.1
12	McKeown, M. Margaret (351)	9	-0.3270	36.9
13	Barksdale, Rhesa H. (675)	5	0.3065	36.6
14	Niemeyer, Paul V. (674)	4	0.5020	35.5
15	Easterbrook, Frank H. (366)	7	0.5590	35.2
16	Murphy, Diana E. (699)	8	-0.5950	34.4
17	Sykes, Diane S. (387)	7	0.5310	34.0
18	Murphy, Michael R. (471)	10	-0.4220	32.1
19	Gilman, Ronald L. (414)	6	-0.4220	29.4
20	Garza, Emilio M. (672)	5	0.5020	27.9
21	Bea, Carlos T. (356)	9	0.5310	27.1
22	Rendell, Marjorie O. (336)	3	-0.4220	25.6
23	Tjoflat, Gerald B. (883)	11	0.4090	25.2
24	Tacha, Deanell R. (390)	10	0.2290	24.0
25	Silverman, Barry G. (338)	9	-0.4220	23.7
26	Gibson, John R. (318)	8	0.1380	23.2
27	Wilkins, William W. (423)	4	0.3900	22.7
28	Wollman, Roger L. (686)	8	0.2180	22.7
29	Edmondson, James L. (379)	11	0.3070	22.5
30	Bauer, William J. (390)	7	0.0340	21.8
31	Wiener, Jacques L., Jr. (739)	5	0.5020	21.2
32	Clifton, Richard R. (304)	9	0.5310	19.9
33	Chagares, Michael A. (375)	3	0.5310	19.4
34	Jones, Edith H. (438)	5	0.5020	19.0
35	Bybee, Jay S. (352)	9	0.5310	18.8
36	Daughtrey, Martha C. (314)	6	-0.2715	18.2
37	Davis, W. Eugene (665)	5	0.5590	18.1
38	Jolly, E. Grady (693)	5	0.2520	18.0
39	Marcus, Stanley (1028)	11	-0.2340	17.2
40	Fay, Peter T. (408)	11	0.4090	17.0
41	Tymkovich, Timothy M. (528)	10	0.3555	16.7
42	Briscoe, Mary B. (504)	10	-0.4220	15.9
43	Baldock, Bobby R. (336)	10	0.2280	15.9
44	Lucero, Carlos F. (447)	10	-0.4220	15.8
45	Hardiman, Thomas M. (374)	3	0.0280	15.0
46	King, Robert B. (735)	4	-0.2605	13.8
47	Manion, Daniel A. (444)	7	0.3515	13.8
48	Smith, Jerry E. (681)	5	0.5020	13.8

Appendix A: Judicial Ideology Scores from the most conservative to most liberal for judges with at least 300 interactions with other judges.

49	Callahan, Consuelo M. (339)	9	0.5310	13.4
50	Flaum, Joel M. (477)	7	0.0340	12.9
51	Shepherd, Bobby E. (618)	8	0.5310	12.2
52	Fuentes, Julio M. (363)	3	-0.3115	11.8
53	Posner, Richard A. (384)	7	0.0340	11.1
54	Kelly, Paul J., Jr. (501)	10	0.2280	10.6
55	Ebel, David M. (381)	10	0.4940	10.4
56	Motz, Diana G. (567)	4	-0.3765	10.3
57	Anderson, Robert L., III (1004)	11	-0.2790	10.2
58	King, Carolyn D. (695)	5	-0.1050	10.0
59	Southwick, Leslie (657)	5	0.3465	9.9
60	Hall, Peter W. (443)	2	0.5310	9.2
61	Cabranes, Jose A. (474)	2	-0.2665	9.2
62	Prado, Edward C. (738)	5	0.3465	8.8
63	Barkett, Rosemary (949)	11	-0.2340	8.0
64	Reavley, Thomas M. (513)	5	-0.1050	7.9
65	Pryor, William H., Jr. (991)	11	0.3955	7.5
66	Shedd, Dennis W. (631)	4	0.3900	7.3
67	Dubina, Joel F. (1020)	11	0.5020	7.3
68	Birch, Stanley F., Jr. (1004)	11	0.5020	7.2
69	Jordan, Kent A. (461)	3	0.5310	6.8
70	Elrod, Jennifer W. (534)	5	0.3465	6.6
71	Hull, Frank M. (1052)	11	-0.1910	6.4
72	Gibbons, Julia S. (368)	6	0.3180	6.2
73	McKee, Theodore A. (336)	3	-0.3080	6.1
74	Black, Susan H. (1042)	11	0.3890	5.7
75	Thomas, Sidney R. (517)	9	-0.2090	5.4
76	Wilkinson, J. Harvie, III (474)	4	0.2335	4.8
77	Howard, Jeffrey R. (346)	1	0.4785	4.6
78	Livingston, Debra A. (421)	2	0.5310	4.4
79	Hamilton, Clyde H. (494)	4	0.3900	4.0
80	Dennis, James L. (697)	5	-0.1070	3.9
81	Hartz, Harris L. (456)	10	0.2280	3.8
82	Evans, Terence T. (489)	7	-0.4280	3.8
83	Traxler, William B., Jr. (714)	4	-0.1910	3.5
84	Clement, Edith B. (691)	5	0.5310	3.5
85	Stewart, Carl E. (758)	5	-0.1070	3.5
86	Carnes, Edward E. (1028)	11	0.5020	3.4
87	Wilson, Charles R. (1112)	11	-0.2340	3.2
88	Benton, William D. (681)	8	0.3010	3.2
89	Ripple, Kenneth F. (414)	7	0.3515	2.6
90	Raggi, Reena (446)	2	0.5310	2.3
91	Colloton, Steven M. (624)	8	0.2760	2.1
92	McKay, Monroe G. (318)	10	-0.5320	0.5
93	Smith, Lavenski R. (635)	8	0.3630	0.0
94	Benavides, Fortunato P. (720)	5	-0.4220	-0.3
95	Bye, Kermit E. (576)	8	-0.2410	-0.4
96	Sack, Robert D. (377)	2	-0.3180	-0.5
97	Sloviter, Dolores K. (360)	3	-0.5320	-0.5
98	Ambro, Thomas L. (402)	3	-0.3110	-0.8
99	Roth, Jane R. (390)	3	0.2370	-0.9
100	Fisher, Raymond C. (305)	9	-0.3380	-1.3

Appendix A: Judicial Ideology Scores from the most conservative to most liberal for judges with at least 300 interactions with other judges.

101	Katzmann, Robert A. (351)	2	-0.3195	-1.3
102	Duncan, Allyson K. (633)	4	0.2950	-1.5
103	Gorsuch, Neil M. (438)	10	0.5310	-1.5
104	Jacobs, Dennis G. (336)	2	0.1350	-3.2
105	Owen, Priscilla R. (662)	5	0.3465	-3.5
106	Anderson, Stephen H. (333)	10	0.4095	-5.2
107	Tashima, A. Wallace (417)	9	-0.3380	-5.5
108	Parker, Barrington D., Jr. (371)	2	0.5310	-5.8
109	Haynes, Catharina (339)	5	0.3465	-6.4
110	Rovner, Ilana D. (426)	7	0.5020	-7.0
111	Sutton, Jeffrey S. (308)	6	0.2260	-7.3
112	Barry, Maryanne T. (386)	3	-0.3115	-7.6
113	Higginbotham, Patrick E. (618)	5	0.4580	-8.4
114	Gregory, Roger L. (665)	4	0.2810	-8.6
115	Wesley, Richard C. (424)	2	0.5310	-8.7
116	Kanne, Michael S. (420)	7	0.3515	-9.4
117	Lipez, Kermit V. (311)	1	-0.4220	-9.5
118	McConnell, Michael W. (453)	10	0.3100	-9.7
119	Kravitch, Phyllis A. (382)	11	-0.0400	-11.8
120	Holmes, Jerome A. (453)	10	0.5725	-12.4
121	Sotomayor, Sonia (360)	2	-0.3180	-14.4
122	Fisher, D. Michael (462)	3	0.1675	-14.6
123	Loken, James B. (354)	8	0.1235	-16.4
124	Lynch, Sandra L. (356)	1	-0.3995	-17.0
125	Martin, Boyce F., Jr. (339)	6	-0.2270	-17.4
126	Williams, Ann C. (384)	7	-0.3450	-18.1
127	Fletcher, Betty B. (647)	9	-0.3400	-19.7
128	Cole, Ransey G., Jr. (417)	6	-0.2670	-20.5
129	Torruella, Juan R. (343)	1	0.5590	-22.4
130	Melloy, Michael J. (453)	8	0.2760	-23.2
131	Pooler, Rosemary S. (357)	2	-0.3180	-24.3
132	Wood, Diane P. (423)	7	-0.3795	-24.9
133	Smith, Norman R. (431)	9	0.4100	-27.8
134	Smith, David B. (375)	3	0.1675	-29.0
135	Paez, Richard A. (416)	9	-0.3380	-29.1
136	Nelson, Dorothy W. (335)	9	-0.4090	-32.2
137	Gould, Ronald M. (329)	9	-0.3270	-34.1
138	Clay, Eric L. (387)	6	-0.3680	-34.4
139	Michael, M. Blane (647)	4	-0.2605	-37.4
140	Moore, Karen N. (465)	6	-0.2670	-37.7
141	Hawkins, Michael D. (346)	9	-0.1650	-38.4
142	Wardlaw, Kim M. (387)	9	-0.3380	-41.7
143	Graber, Susan (356)	9	-0.2920	-66.7

Appendix B: Judicial Ideology Scores by circuit for judges with at least 300 interactions with other judges.

Rank	Judge (Total Interactions)	Circuit	Common Space Score	Ideology Score
77	Howard, Jeffrey R. (346)	1	0.4785	4.6
117	Lipez, Kermit V. (311)	1	-0.422	-9.5
124	Lynch, Sandra L. (356)	1	-0.3995	-17.0
129	Torruella, Juan R. (343)	1	0.559	-22.4
61	Cabranes, Jose A. (474)	2	-0.2665	9.2
60	Hall, Peter W. (443)	2	0.531	9.2
104	Jacobs, Dennis G. (336)	2	0.135	-3.2
101	Katzmann, Robert A. (351)	2	-0.3195	-1.3
78	Livingston, Debra A. (421)	2	0.531	4.4
108	Parker, Barrington D., Jr. (371)	2	0.531	-5.8
131	Pooler, Rosemary S. (357)	2	-0.318	-24.3
90	Raggi, Reena (446)	2	0.531	2.3
96	Sack, Robert D. (377)	2	-0.318	-0.5
121	Sotomayor, Sonia (360)	2	-0.318	-14.4
115	Wesley, Richard C. (424)	2	0.531	-8.7
98	Ambro, Thomas L. (402)	3	-0.311	-0.8
112	Barry, Maryanne T. (386)	3	-0.3115	-7.6
33	Chagares, Michael A. (375)	3	0.531	19.4
122	Fisher, D. Michael (462)	3	0.1675	-14.6
52	Fuentes, Julio M. (363)	3	-0.3115	11.8
45	Hardiman, Thomas M. (374)	3	0.028	15.0
69	Jordan, Kent A. (461)	3	0.531	6.8
73	McKee, Theodore A. (336)	3	-0.308	6.1
22	Rendell, Marjorie O. (336)	3	-0.422	25.6
99	Roth, Jane R. (390)	3	0.237	-0.9
97	Sloviter, Dolores K. (360)	3	-0.532	-0.5
134	Smith, David B. (375)	3	0.1675	-29.0
102	Duncan, Allyson K. (633)	4	0.295	-1.5
114	Gregory, Roger L. (665)	4	0.281	-8.6
79	Hamilton, Clyde H. (494)	4	0.39	4.0
46	King, Robert B. (735)	4	-0.2605	13.8
139	Michael, M. Blane (647)	4	-0.2605	-37.4
56	Motz, Diana G. (567)	4	-0.3765	10.3
14	Niemeyer, Paul V. (674)	4	0.502	35.5
66	Shedd, Dennis W. (631)	4	0.39	7.3
83	Traxler, William B., Jr. (714)	4	-0.191	3.5
27	Wilkins, William W. (423)	4	0.39	22.7
76	Wilkinson, J. Harvie, III (474)	4	0.2335	4.8
13	Barksdale, Rhesa H. (675)	5	0.3065	36.6
94	Benavides, Fortunato P. (720)	5	-0.422	-0.3
84	Clement, Edith B. (691)	5	0.531	3.5
37	Davis, W. Eugene (665)	5	0.559	18.1
80	Dennis, James L. (697)	5	-0.107	3.9
70	Elrod, Jennifer W. (534)	5	0.3465	6.6
20	Garza, Emilio M. (672)	5	0.502	27.9

Appendix B: Judicial Ideology Scores by circuit for judges with at least 300 interactions with other judges.

109	Haynes, Catharina (339)	5	0.3465	-6.4
113	Higginbotham, Patrick E. (618)	5	0.458	-8.4
38	Jolly, E. Grady (693)	5	0.252	18.0
34	Jones, Edith H. (438)	5	0.502	19.0
58	King, Carolyn D. (695)	5	-0.105	10.0
105	Owen, Priscilla R. (662)	5	0.3465	-3.5
62	Prado, Edward C. (738)	5	0.3465	8.8
64	Reavley, Thomas M. (513)	5	-0.105	7.9
48	Smith, Jerry E. (681)	5	0.502	13.8
59	Southwick, Leslie (657)	5	0.3465	9.9
85	Stewart, Carl E. (758)	5	-0.107	3.5
31	Wiener, Jacques L., Jr. (739)	5	0.502	21.2
3	Batchelder, Alice M. (308)	6	0.502	80.4
138	Clay, Eric L. (387)	6	-0.368	-34.4
128	Cole, Ransey G., Jr. (417)	6	-0.267	-20.5
36	Daughtrey, Martha C. (314)	6	-0.2715	18.2
72	Gibbons, Julia S. (368)	6	0.318	6.2
19	Gilman, Ronald L. (414)	6	-0.422	29.4
5	Griffin, Richard A. (421)	6	0.531	61.2
125	Martin, Boyce F., Jr. (339)	6	-0.227	-17.4
2	McKeague, David W. (386)	6	0.531	91.3
140	Moore, Karen N. (465)	6	-0.267	-37.7
1	Rogers, John M. (366)	6	0.392	96.4
11	Siler, Eugene E., Jr. (325)	6	0.339	43.1
111	Sutton, Jeffrey S. (308)	6	0.226	-7.3
30	Bauer, William J. (390)	7	0.034	21.8
15	Easterbrook, Frank H. (366)	7	0.559	35.2
82	Evans, Terence T. (489)	7	-0.428	3.8
50	Flaum, Joel M. (477)	7	0.034	12.9
116	Kanne, Michael S. (420)	7	0.3515	-9.4
47	Manion, Daniel A. (444)	7	0.3515	13.8
53	Posner, Richard A. (384)	7	0.034	11.1
89	Ripple, Kenneth F. (414)	7	0.3515	2.6
110	Rovner, Ilana D. (426)	7	0.502	-7.0
17	Sykes, Diane S. (387)	7	0.531	34.0
126	Williams, Ann C. (384)	7	-0.345	-18.1
132	Wood, Diane P. (423)	7	-0.3795	-24.9
88	Benton, William D. (681)	8	0.301	3.2
95	Bye, Kermit E. (576)	8	-0.241	-0.4
91	Colloton, Steven M. (624)	8	0.276	2.1
26	Gibson, John R. (318)	8	0.138	23.2
6	Gruender, Raymond W. (602)	8	0.301	56.9
123	Loken, James B. (354)	8	0.1235	-16.4
130	Melloy, Michael J. (453)	8	0.276	-23.2
16	Murphy, Diana E. (699)	8	-0.595	34.4
8	Riley, William J. (573)	8	0.288	49.7
51	Shepherd, Bobby E. (618)	8	0.531	12.2
93	Smith, Lavenski R. (635)	8	0.363	0.0
28	Wollman, Roger L. (686)	8	0.218	22.7

Appendix B: Judicial Ideology Scores by circuit for judges with at least 300 interactions with other judges.

21	Bea, Carlos T. (356)	9	0.531	27.1
35	Bybee, Jay S. (352)	9	0.531	18.8
49	Callahan, Consuelo M. (339)	9	0.531	13.4
10	Canby, William C., Jr. (320)	9	-0.165	46.7
32	Clifton, Richard R. (304)	9	0.531	19.9
100	Fisher, Raymond C. (305)	9	-0.338	-1.3
127	Fletcher, Betty B. (647)	9	-0.34	-19.7
137	Gould, Ronald M. (329)	9	-0.327	-34.1
143	Graber, Susan (356)	9	-0.292	-66.7
141	Hawkins, Michael D. (346)	9	-0.165	-38.4
9	Ikuta, Sandra S. (329)	9	0.531	49.0
12	McKeown, M. Margaret (351)	9	-0.327	36.9
136	Nelson, Dorothy W. (335)	9	-0.409	-32.2
4	O'Scannlain, Diarmuid F. (417)	9	0.023	72.6
135	Paez, Richard A. (416)	9	-0.338	-29.1
25	Silverman, Barry G. (338)	9	-0.422	23.7
133	Smith, Norman R. (431)	9	0.41	-27.8
107	Tashima, A. Wallace (417)	9	-0.338	-5.5
75	Thomas, Sidney R. (517)	9	-0.209	5.4
142	Wardlaw, Kim M. (387)	9	-0.338	-41.7
106	Anderson, Stephen H. (333)	10	0.4095	-5.2
43	Baldock, Bobby R. (336)	10	0.228	15.9
42	Briscoe, Mary B. (504)	10	-0.422	15.9
55	Ebel, David M. (381)	10	0.494	10.4
103	Gorsuch, Neil M. (438)	10	0.531	-1.5
81	Hartz, Harris L. (456)	10	0.228	3.8
120	Holmes, Jerome A. (453)	10	0.5725	-12.4
54	Kelly, Paul J., Jr. (501)	10	0.228	10.6
44	Lucero, Carlos F. (447)	10	-0.422	15.8
118	McConnell, Michael W. (453)	10	0.31	-9.7
92	McKay, Monroe G. (318)	10	-0.532	0.5
18	Murphy, Michael R. (471)	10	-0.422	32.1
7	O'Brien, Terrence L. (345)	10	0.4405	56.1
24	Tacha, Deanell R. (390)	10	0.229	24.0
41	Tymkovich, Timothy M. (528)	10	0.3555	16.7
57	Anderson, Robert L., III (1004)	11	-0.279	10.2
63	Barkett, Rosemary (949)	11	-0.234	8.0
68	Birch, Stanley F., Jr. (1004)	11	0.502	7.2
74	Black, Susan H. (1042)	11	0.389	5.7
86	Carnes, Edward E. (1028)	11	0.502	3.4
67	Dubina, Joel F. (1020)	11	0.502	7.3
29	Edmondson, James L. (379)	11	0.307	22.5
40	Fay, Peter T. (408)	11	0.409	17.0
71	Hull, Frank M. (1052)	11	-0.191	6.4
119	Kravitch, Phyllis A. (382)	11	-0.04	-11.8
39	Marcus, Stanley (1028)	11	-0.234	17.2
65	Pryor, William H., Jr. (991)	11	0.3955	7.5
23	Tjoflat, Gerald B. (883)	11	0.409	25.2
87	Wilson, Charles R. (1112)	11	-0.234	3.2