ARTICLES

CHOOSING THE NEXT SUPREME COURT JUSTICE: AN EMPIRICAL RANKING OF JUDGE PERFORMANCE

STEPHEN J. CHOI* & G. MITU GULATI**

ABSTRACT

The judicial appointments process has grown increasingly frustrating in recent years. Both sides claim that their candidates are the "most

---

* Roger J. Traynor Professor, U.C. Berkeley Law School (Boalt Hall).

** Professor of Law, Georgetown University. Kindly e-mail comments to choi@law.berkeley.edu and gulati@law.georgetown.edu. Erin Dengan, Jennifer Dukart, Édeanna Johnson-Chebbi, Alice Kuo, Margaret Rodgers, and Rishi Sharma provided research assistance. Kimberly Brickell deserves special thanks for her work. Aspects of this draft benefited from discussions with Alex Aleinikoff, Ian Ayres, Scott Baker, Barbara Banoff, Devon Carbado, Lee Epstein, Tracey George, Prea Gulati, Vicki Jackson, Mike Klarman, Kim Krawiec, Paul Mahoney, Kaleb Michaud, Greg Mitchell, Ed Kitch, Un Kyung Park, Jim Rossi, Jim Ryan, Mark Seidenfeld, Howard Shelanski, Paul Stefan, Eric Talley, and George Triantis. For comments on the draft itself, we are grateful to Michael Bailey, Suzette Baker, Bill Bratton, James Brudney, Brannon Denning, Dan Farber, Phil Frickey, Michael Gerhardt, Steve Goldberg, Linc Kaplan, Joe Kennedy, Vikram Khanna, Pauline Kim, Don Langevoort, Jim Lindgren, Bill Marshall, Steve McBundy, Eric Mueller, Jeff Rachlinksi, Judith Resnik, Steve Salop, Michael Seidman, Peter Seigelman, Keith Sharfman, Pat Shin, Michael Solimine, Larry Solum, Gerry Spann, Nancy Staudt, Mark Tushnet, David Vladeck, David Walker, Robin West, Kathy Zeiler, Arnold Zellner, Todd Zywicki, and participants at workshops at Berkeley, Boston University, Florida State, Georgetown, Virginia, and UNC-Chapel Hill. Given the unusually large number of people who have e-mailed us with comments on this project, it is likely that there are some who we have inadvertently failed to thank. Our sincerest apologies to them.

Disclosure: Funding for this project was provided entirely by our respective law schools. One of us was a law clerk to two of the judges in the sample: Samuel Alito of the Third Circuit and Sandra Lynch of the First Circuit. The individuals and institutions mentioned above bear no responsibility for the ideas expressed here.

meritorious" and yet there is seldom any discussion of what constitutes merit. Instead, the discussion moves immediately to the candidates' likely positions on hot-button political issues like abortion, gun control, and the death penalty. One side claims that it is proposing certain candidates based on merit, while the other claims that the real reason for pushing those candidates is their ideology and, in particular, their likely votes on key hot-button issues. With one side arguing merit and the other side arguing ideology, the two sides talk past each other and the end result is often an impasse. To get past this impasse, we propose placing judges in a tournament based on relatively objective measures of judicial merit and productivity. A tournament allows the public to test the politicians' claims of merit. Being able to test these claims helps make transparent the occasions in which the real debate is over ideology. It is harder to disguise a purely ideological candidate as the best from a "merit" standpoint when the candidate performs poorly relative to many other judges based on objective factors. Once merit-based arguments have been isolated (or at least reduced in scope) to factors related to the tournament, it should be possible to have a transparent and meaningful debate over ideology.

This Article runs such a tournament using data on opinions authored by active federal circuit court judges from one common time period: the beginning of 1998 to the end of 2000. The focus on a common time period helps put judges in the tournament on a level playing field. We then generate a series of measures of merit focusing on (a) productivity, (b) opinion quality, and (c) judicial independence. While not perfect, our measures interject a greater focus on merit in the current nomination process (thereby revealing previously nontransparent motives based on ideology). With our data, we are able to test the claims of merit that the next president will inevitably make when he announces one of his favorite circuit court judges as the nominee for the Supreme Court.

TABLE OF CONTENTS

I. INTRODUCTION: CAN ONE MEASURE JUDICIAL MERIT?

In the next couple of years the Supreme Court will likely have one to three seats fall vacant. Justices Stevens, Rehnquist, and O'Connor are all either in their eighties or approaching that point, and rumors suggest that at least two of them are contemplating retirement.[1] Washington, D.C. is abuzz with speculation as to who might be the favored candidates. Once the president settles on a nominee, the following scenario will unfold: he will introduce his candidate with something along the lines of "Having been a distinguished circuit court judge for many years, Judge Y is highly qualified for the position of Associate Justice of the United States Supreme Court." That announcement will spark a frenzy of inquiry into the candidate's past. The focus will be on the candidate's expected position on hot-button political issues such as abortion, the death penalty, and affirmative action. Discussion of the candidate's broader qualifications will be pushed to the background. At best, the press accounts will carry a brief mention of the candidate's employment history and the name of the law school the candidate attended. There may be a quote or two from a former colleague or classmate, but there will be little in the way of systematic analysis of the candidate's past performance. This includes that which the

1. *See, e.g.*, Sheryl Gay Stolberg, *The War over Abortion Moves to a Smaller Stage*, N.Y. TIMES, Oct. 26, 2003, § 4, at 4 ("Both sides [of the abortion debate] know that with one or more Supreme Court justices now contemplating retirement, [the Court] could easily flip the other way if Republicans keep their hold on the White House."). On the ages of the current Justices and speculation about potential nominees, see Kenneth L. Manning, Bruce A. Carroll & Robert A. Carp, *George W. Bush's Potential Supreme Court Nominees: What Impact Might They Have?*, 85 JUDICATURE 278, 278–79 (2001–2002). As this Article entered its final editing stage, the media reported Chief Justice Rehnquist's thyroid cancer, increasing even more the speculation over who the next nominee will be. *See, e.g.*, Robert S. Greenberger, *Rehnquist Note Suggests Cancer Is Severe Variety*, WALL ST. J., Nov. 2, 2004, at A2.

president touted at the outset: *the candidate's career as a circuit court judge*.[2]

The genesis of this project lies in our frustration with the current appointment process. As best as we can tell, the entire focus in analyzing a candidate's qualifications is on predicting their expected votes on a handful of issues. And politicians can tout their respective favorite candidates as "highly qualified" and "intellectually superior" with little challenge.[3] In a number of recent high-profile nominations to the federal courts of appeals—including Miguel Estrada, Justice Priscilla Owen, and Judge William Pryor—the following pattern has played out. The Republicans first tout their nominee as "highly qualified" and most deserving of confirmation. The Democrats then mount an opposition on the grounds that the nominee is too ideological. The first side then responds by saying that its opponents are ideologues because they are blocking a highly qualified candidate.

For purposes of this Article, we make two observations regarding the above scenario. First, both sides seem to perceive that their constituents want selections to be based on merit and not political ideology.[4] Hence, even if the politicians care not a whit about merit and only care about ideology, they are constrained by the need to justify their selections as

---

2. In recent years, a norm appears to have developed in which experience on a federal circuit court (with special importance given to the D.C. Circuit) is seen as a key qualification for elevation to the Court. *See* Lee Epstein, Jack Knight & Andrew D. Martin, *The Norm of Prior Judicial Experience and Its Consequences for Career Diversity on the U.S. Supreme Court*, 91 CAL. L. REV. 903 (2003).

3. *See, e.g.*, Helen Dewar, *GOP Presses for Votes on Judges; Senate Republicans Force New Vote on One Nominee, but Democrats Vow to Prevail*, WASH. POST, July 30, 2003, at A4 (observing that the Republicans describe Priscilla Owen as "highly qualified," whereas the Democrats see her as a "pro-business, anti-abortion activist who lets her personal beliefs guide her legal actions"), *available at* 2003 WL 56509262; Orrin Hatch, *Abortion Stances Based in Religion*, ROLL CALL, Sept. 8, 2003, at Letters section (claiming that William Pryor, President Bush's "highly qualified" judicial nominee, was being opposed on religious grounds), *available at* 2003 WL 7691833; Neil A. Lewis, *GOP Senators Try to Change Filibuster Rules*, SAN DIEGO UNION TRIB., May 10, 2003, at A6 (quoting Senate Republican leader Bill Frist as referring to both Miguel Estrada and Justice Priscilla Owen as "highly qualified and intellectually superior"), *available at* 2003 WL 6582851; James Wensits, *Chocola Supports Bush Court Nominee*, S. BEND TRIB., Feb. 14, 2003, at D3 (quoting U.S. Representative Chris Chocola as asserting, "There is no question that Miguel Estrada is highly qualified to serve on the federal bench," without providing substantial evidence beyond Estrada's schooling and his employment in the Justice Department), *available at* 2003 WL 9896880; *President's Statement on the Senate Filibuster of Judicial Nominees*, 39 WEEKLY COMPILATION OF PRESIDENTIAL DOCUMENTS 1025, 1025–26 (2003) (claiming that President Bush's "highly qualified nominees" with "stellar records" are being blocked), *available at* 2003 WL 13973558.

4. For example, President Bill Clinton asserted during his 1992 campaign that, "if given the opportunity," he would select Justices of "'far higher quality'" than his Republican predecessors and that he would "'take [the process] out of politics.'" JOHN ANTHONY MALTESE, THE SELLING OF SUPREME COURT NOMINEES 150 (1995) (quoting President Clinton).

meritorious. Second, we suspect there is little quarrel over the merit of the nominees because politicians care little about it and the public has no means of evaluating it.[5] Instead, claims of merit are largely ignored and the quarrel takes place over ideology, with each side claiming that ideology is driving the behavior of their opponents in either proposing or opposing a nominee. The problem, from the public's perspective, is that it is difficult to determine what is going on. That is, when is a highly qualified candidate being blocked unfairly? Or, when is a highly ideological candidate being blocked fairly?

Of course, what one person calls ideology may be what another calls merit.[6] Someone in favor of affirmative action, for example, may deem a judicial nominee who supports affirmative action as meritorious. We have a slightly different set of meanings in mind. While different visions of merit may exist, some are more widely held than others. Few would quarrel with the claim that a judge who displays productivity, intelligence, and integrity is better than one who does not. On the other hand, less broad-based support exists for judges favoring affirmative action. Our use of "merit" refers to more widely held views of what makes a good judge. "Ideology," on the other hand, is used to refer to narrowly held views. The problem we address is the tendency for politicians to mask their support for a particular judge based on more narrowly held ideology with generic and noninformative references to widely held criteria masking the true reasons for their support for a particular candidate. Our goal with the tournament is to force politicians to come clean on their motivations. Confronting politicians with a set of objective measures of merit, we hope, will force them to explain how they can claim a candidate is the most qualified if she does not do well on our measures.

The rhetoric about merit aside, and assuming that merit and ideology are separable, there is a long history of arguments that politics and ideology

5. Bill Marshall and Michael Gerhardt, both of whom have first-hand experience with the appointments process, tell us that we have overstated our point about politicians' lack of concern about merit. To the extent that politicians have political agendas that they (or interest groups) are seeking to push via certain judicial nominees, judges who are more capable at their job will be better able to push the agendas than those who are not as capable. From this standpoint, merit matters to the politicians. It may, therefore, be more accurate to say that politicians are unconcerned about *excellence*, independent of ideology, rather than *merit*.

6. As with merit, it is difficult to define ideology. For a discussion of the definition of ideology, see Stephen B. Presser, *Should Ideology of Judicial Nominees Matter?: Is the Senate's Current Reconsideration of the Confirmation Process Justified?*, 6 TEX. REV. L. & POL. 245, 246–47 (2001) (drawing a distinction between an ideology of process and one of substance and arguing that it is the former that is legitimate).

should matter in the nomination process.[7] If politics motivate the president's choice or the Senate's resistance, the tournament helps make this factor more transparent. Rather than hide behind "most qualified," the president and the Senate must make more explicit the political litmus test used to justify the candidate. So, for example, if there is a clear ranking of the qualifications of the over 160 active circuit court judges and the president chooses to nominate the judge ranked number forty-two to the Supreme Court, as opposed to one of the top rankers, that should raise suspicion about the claim that the nominee is the most qualified. Conversely, the claim that ideology is driving the president's selection begins to seem more plausible. To the extent that the public concludes that one side is being unduly ideological, it can penalize it in the next elections.

Consider specific types of merit claims that are made about candidates and how objective measures of performance can be used to challenge such claims. Politicians will sometimes directly invoke the criteria we attempt to measure: independence, effort, and quality. For example, we found numerous instances in which politicians touted particular candidates because of their independent thinking and reputations, usually in response to an attack that claimed that the candidate was ideologically driven.[8] If a

---

7. *See, e.g.*, Erwin Chemerinsky, *Ideology and the Selection of Federal Judges*, 36 U.C. DAVIS L. REV. 619, 627–31 (2003). For more on the arguments that ideology should play or has played an important role in the process, see John S. Baker, Jr., *Ideology and the Confirmation of Federal Judges*, 43 S. TEX. L. REV. 177 (2001); James J. Brudney, *Recalibrating Federal Judicial Independence*, 64 OHIO ST. L.J. 149, 153–61 (2003); John C. Eastman, *The Limited Nature of the Senate's Advice and Consent Role*, 36 U.C. DAVIS L. REV. 633, 648–49 (2003) (recalling that ideology was of primary concern with President George Washington's nominee, John Rutledge); Ed R. Haden, *Judicial Selection: A Pragmatic Approach*, 24 HARV. J.L. & PUB. POL'Y 531 (2001); William G. Ross, *The Role of Judicial Issues in Presidential Campaigns*, 42 SANTA CLARA L. REV. 391 (2002); Matthew D. Marcotte, Note, *Advice and Consent: A Historical Argument for Substantive Senatorial Involvement in Judicial Nominations*, 5 N.Y.U. J. LEGIS. & PUB. POL'Y 519 (2001–2002).

For an argument that ideology and politics should not matter in the judicial selection process, see Ronald D. Rotunda, *The Role of Ideology in Confirming Federal Court Judges*, 15 GEO. J. LEGAL ETHICS 127, 131 (2001) (asserting that nominees should not be asked specific legal questions because the purpose of the confirmation is to ensure a fair and impartial court, not one with political leanings).

8. For example, Senator Warren Rudman praised Justice Souter during the nomination process by asserting that "'[h]e can't be classified as an ideologue in any way, shape or fashion'" and "'does not have an agenda of his own.'" Ruth Marcus & David S. Broder, *Bush Names Appellate Judge to Brennan Seat; Conservative's Slate Blank on Abortion*, WASH. POST, July 24, 1990, at A1 (quoting Senator Rudman), *available at* 1990 WL 2119014. Yet, there is a long history of prospective candidates for the Court being attacked as not being independent. *See, e.g.*, MALTESE, *supra* note 5, at 41, 50–51, 54–55 (describing how the question of independence came up in the nominations of Stanley Matthews, Louis D. Brandeis, and Charles Evans Hughes).

One can see the invocation of independence as a criterion at the circuit court level as well. For example, Senator Hatch said that Michael McConnell (a nominee for a seat on the Tenth Circuit) "cannot be pegged as an ideologue in any sense" and "is beholden to no one or no group." *Nomination of Michael W. McConnell to Tenth Circuit Before the Senate Judiciary Comm.*, 107th Cong. (2002)

politician were to make a claim based on independence about a candidate, and the data showed that the judge had always voted in line with those who shared the politician's views and only voted against those with different views, the claim of independence would look problematic. Similarly, if a politician were to claim that a candidate was especially qualified because of the intellect and scholarly ability that the candidate had demonstrated as a jurist, the fact that the candidate's opinions were never cited by other judges would make that claim look problematic.[9]

This Article presents a set of simple and objective measures to evaluate judicial merit, placing judges in a tournament of sorts using criteria correlated (albeit imperfectly) with widely held notions of merit. Our simple measures do not provide a perfect metric for judging skill, but that is not the standard at which we are aiming. The goal is to demonstrate

---

(statement of Sen. Orrin Hatch), *available at* http://hatch.senate.gov/index.cfm?FusionAction=PressReleases.Detail&PressRelease_id=573&Month=9&Year=2002. Hatch also asserted that a Ninth Circuit nominee, Judge Carolyn Kuhl, was not biased and quoted letters of support from her colleagues at the Litigation Section of the Los Angeles County Bar Association and the Los Angeles County Superior Court saying, respectively, that she is fair-minded and "administers justice without favor." *Hearing on Nomination of Carolyn Kuhl for Ninth Circuit Before the Senate Judiciary Comm.*, 108th Cong. (2003), (statement of Sen. Orrin Hatch), *available at* http://hatch.senate.gov/index.cfm?FusionAction=PressReleases.Detail&PressRelease_id=739&Month=4&Year=2003. During the Priscilla Owen nomination process, Senator Hatch pointed to the fact that "she is not afraid to dissent" (presumably a claim about her independence). *Nomination of Priscilla Owen, U.S. Circuit Judge, Fifth Circuit Before the Senate Judiciary Comm.*, 107th Cong. (2002) (statement of Sen. Orrin Hatch), *available at* http://hatch.senate.gov/index.cfm?FusionAction=PressReleases.Detail&PressRelease_id=542&Month=7&Year=2002.

9. To reference Justice Souter again, claims about his scholarly ability and intellect were explicitly made during his nomination process. Referring to Souter's background as an associate justice on the New Hampshire Supreme Court and (ironically, given his short tenure there) his experience on the First Circuit Court of Appeals, President George H.W. Bush described him as "a remarkable judge of keen intellect and the highest ability, one whose scholarly commitment to the law and whose wealth of experience mark him of first rank." WASH. POST, July 24, 1990, at A12, *available at* 1990 WL 2118996. Bush also said, "His opinions reflect a clean intellect-keen intellect, as well as wise balance between the theoretical and practical aspects of the law." *Id.* Senator Rudman described Souter as "'a brilliant intellectual, a classic conservative intellectual in the deepest sense of the word.'" Marcus & Broder, *supra* note 8. Deputy Attorney General William P. Barr said, "'I think the attraction [to Souter] was that invariably his opinions appeared to be very scholarly, ably written and he appeared to be a believer in judicial restraint.'" Ann Devroy, *Bush Names Appellate Judge to Brennan Seat; President Selects Souter, 50, for 'Intellect' and 'Ability,'* WASH. POST, July 24, 1990, at A1, *available at* 1990 WL 2119016.

Although our sense is that attacks on candidates for having mediocre intellect have been rare, they have occurred. *See* MALTESE, *supra* note 4, at 16 (describing how G. Harrold Carswell was criticized as inept). At least one senator, though, Senator Roman Hruska, famously defended Carswell's mediocrity. He said, "Even if he is mediocre, there are a lot of mediocre judges and people and lawyers. They are entitled to a little representation, aren't they? We can't have all Brandeises, Cardozos and Frankfurters, and stuff like that there." *Id.* at 16. Hruska's defense of Carswell's mediocrity left the White House in shock. *Id.*

the availability of a set of objective measures for which we can easily collect data and analyze and that would better identify, at the outset, a merit-worthy pool of Supreme Court candidates. We suspect, under the current system, that merit is used to disguise less than merit-worthy political motivations. At the least, our proposed introduction of a norm to apply objective criteria will force politicians to provide more justification for their selection.

Some will see the search for a set of objective measures as pointless because they think that there is no way to measure or quantify what it means to be a good, let alone great, judge. This is likely true as an *absolute* matter. Nonetheless, with a set of candidates with track records as lower court judges, it may still be possible to make meaningful *relative* evaluations. So, just as it is impossible to articulate what special factor makes Lance Armstrong the best cyclist in the world, it is impossible to reduce Justice Benjamin Cardozo's greatness as a judge to numbers. But one can look at how many times Armstrong has won the *Tour de France* and compare his numbers to those of his peers. Similarly, one can look at Justice Cardozo's opinions and see how often they were cited by other judges, how often they were discussed in law reviews, and how often they made their way into casebooks. Justice Cardozo's numbers can then be compared to those of his peers.[10] As with Armstrong, this type of relative analysis does not give us a measure of his greatness or tell us what made him great. But it gives us a sense, even if imperfect, of how he performed relative to his peers.

Even with the possibility of relative evaluations, no one set of objective measures may obtain popular or even academic consensus. Nonetheless, our approach launches a discussion on how to develop a widely accepted set of objective criteria. Objective criteria may never help us select the very "best" candidate for promotion to the Supreme Court, but

---

10. Richard Posner, in his work on Justice Cardozo and Judge Learned Hand, looked at a set of such measures. *See* RICHARD A. POSNER, CARDOZO: A STUDY IN REPUTATION 74–91 (1990); Richard A. Posner, *The Learned Hand Biography and the Question of Judicial Greatness*, 104 YALE L.J. 511, 534–40 (1994) (book review).

Some of our critics assert that Justice Cardozo makes a bad example of relative excellence because the view of his greatness held by many derives from his performance at the state court level and not on the Supreme Court (where he had only a six-year stint). In starker terms, the argument is that the fact that Justice Cardozo had high citation numbers on the state court level but then had a mediocre career on the Supreme Court goes to show that citations are a flawed predictor of Supreme Court performance. Note, however, that a survey of eight of the best known lists of "great" Justices reveals that Justice Cardozo makes it onto seven of them. *See* Lee Epstein, Tracey George, Micheal Giles & Thomas G. Walker, Rating the Justices: Lessons from Another Court 2 (April 1992) (unpublished draft, presented at the Midwest Political Science Association Meeting, on file with authors).

if the process narrows the field of potential candidates to a substantial extent—say, to the top ten judges selected with a variety of different objective criteria relative to their peers—then putting circuit judges in a tournament to determine Supreme Court nominees will have a significant effect on the nomination process. Rather than simply pronouncing a candidate as the most qualified, the president will face pressure to either select from one of the tournament winners or, in the alternative, to justify why the candidate is the best despite failing to perform well in the tournament. If the president were to choose one of the low-performing candidates, it would suggest that the candidate did the best on the president's own set of political litmus tests.

So how should we go about designing objective criteria of judicial merit? Most, if not all, candidates for the Court possess track records as judges. Since the question is whether the candidate should be promoted to the Supreme Court, step one should be to evaluate the candidate's performance as a lower-level judge. After all, the job at the lower court level of deciding cases and writing opinions that explain the decisions is often much the same at the higher level, the difference being that one hears fewer, but more important, cases. Past performance is key for two reasons. One, assuming the two jobs are similar enough, performance at the lower level helps predict performance at the higher level.[11] Two, the knowledge that the best performances at the lower level will be rewarded with a promotion helps motivate the lower-level employees to exert greater effort. This is the rationale for promotion tournaments in almost every employment setting. Why should this logic not work with judges? In a prior piece, we asked this question and attempted to answer objections.[12] This Article takes on the harder question of how one might implement such a tournament and then does so for the set of active federal circuit judges with published opinions for the three year period from 1998 to 2000.

---

11. A critic might question the link that we make between the job of a circuit judge and that of a Supreme Court Justice. Certainly differences exist. Because the Court is the final arbiter on most issues and is not bound by precedent to the degree that the lower courts are bound, Supreme Court decisionmaking involves more policymaking than the lower courts. In a similar vein, because there are nine Justices and because of the importance of anything that the Court says, Supreme Court decisionmaking involves a greater need for negotiation and compromise than on a circuit court, where the issues are often mundane. Additionally, a significant portion of a Supreme Court Justice's job involves deciding which cases to deny certiorari, something that lower court judges do not do at all. The key question, however, is whether the current system of using litmus tests on issues such as abortion does any better than our tournament in predicting things like the ability to compromise and negotiate, policymaking abilities, or certiorari-denying skills. We do not see how it could.

12. *See* Stephen Choi & Mitu Gulati, *A Tournament of Judges?*, 92 CAL. L. REV. 299 (2004).

Others have studied citation counts and more generally the question of which circuit court judge is the best.[13] Our study differs along at least two dimensions. First, because we are interested in finding the best nominee today, the question we ask is not the more general one of who is the best or most influential over their entire career, but rather who is the best today and for the near future. This allows us to focus on a relatively contemporary and common time period, from 1998 to 2000, to assess the various circuit court judges. Focusing on the present leads to some interesting results. Judge Richard Posner, who uniformly does well in prior citation studies, also does well in our study. He is no longer universally the best, however. For one of our measures of opinion quality, the number of outside-circuit citations garnered by a judge's top twenty citation-receiving opinions, Judge Sandra Lynch of the First Circuit is the highest ranking judge. Second, we expand on prior studies through the introduction of a new measure of merit: independence. In addition to how productive judges are and how often they are cited, we are concerned with how independent a thinker a particular judge is.

Part II compares the current system to our tournament and uses that comparison as a basis to evaluate the proposal's likelihood of adoption. Part III describes the basic building blocks of the tournament. Using data from opinions authored during the 1998 to 2000 time period, Part III reports how active federal circuit court judges fared relative to one another along a number of criteria. Part IV examines how to combine the criteria into a composite metric and reports the winner(s) of the tournament. Part V focuses on five present circuit court judges rumored to be on President Bush's short list of potential Supreme Court nominees.

II. THE CURRENT SYSTEM COMPARED

A promotion tournament of sorts already exists for Supreme Court Justice nominees. Implicit criteria for nomination appear to include candidates who are federal circuit court judges, candidates who are not too old (the older a candidate is, the shorter the candidate's period of influence on the Court), and candidates appointed to the bench by a president of the

13. *See, e.g.*, William M. Landes, Lawrence Lessig & Michael E. Solimine, *Judicial Influence: A Citation Analysis of Federal Courts of Appeals Judges*, 27 J. LEGAL STUD. 271 (1998) (studying the influence of individual circuit court judges with at least six years of tenure by 1995, based on citation counts for the opinions written by each judge). *Cf.* Richard A. Posner, *Is the Ninth Circuit Too Large? A Statistical Study of Judicial Quality*, 29 J. LEGAL STUD. 711 (2000) (providing an empirical test of whether the quality of Ninth Circuit opinions is harmed by the large number of judges on the circuit based on Supreme Court reversals and citation counts).

same political party as the current president. For the sake of simplicity, we use these criteria to constitute the pool of available candidates. Additionally, our quarrel is not with this first step; it is with the next step of the current process.

In step two of the current process, the president seems to narrow the candidate pool on the basis of likely votes on a key subset of political issues such as abortion, gay rights, affirmative action, sexual harassment, the death penalty, gun control, and federalism. The candidate's likely votes on this subset of key issues become a proxy for the nominee's fuller range of future voting behavior. And, as we know from newspaper reports of the recent fights over judicial nominations, the candidate's judging record and personal life are magnified and scrutinized to discern all possible signals of future voting patterns.[14]

This process is flawed. A Supreme Court Justice decides cases on a much broader range of topics than the politically charged issues. What should be a reasonable debate on a candidate's ideology, such as whether the candidate is a strict constructionist, is reduced to quibbling over the candidate's expected position on issues like affirmative action and abortion. That, in turn, means there is room for strategic judging. Lower court judges with aspirations to be on the High Court have an incentive to vote on the political issues in ways that will most please their potential sponsors (and, if possible, least offend their potential opponents). Thus, judges with higher aspirations need not be productive—or move their docket along to resolve disputes promptly and efficiently—nor need they draft well-written opinions oft cited by their colleagues. A judge need only vote the party line. The consequences of improving on the current system are therefore broader than simply inspiring judges, who are appointed for life, not to politicize their opinions or to shirk their basic obligations to decide cases and explain their reasoning. The current selection process and its focus on highly political proxies of a candidate's ideology compromises the independence and productivity of the entire judiciary.

Take, by contrast, our tournament. Like the current system, we also propose a set of proxies. But rather than focusing on the candidate's expected votes on hot-button political issues, we propose to predict (and reward) the candidates' performances on three relatively apolitical factors: (a) productivity, (b) quality, and (c) independence. People will disagree with our methods of measuring performance because they are imperfect. We use a set of objective measures such as number of opinions written (as

14. *See, e.g.*, Dewar, *supra* note 3; Wensits, *supra* note 3.

a proxy for productivity) and number of citations (as a proxy for quality). Citations are a highly imperfect measure of quality in the same way that the SAT is a highly imperfect measure of aptitude for college; so, like the SAT, our measure may have embedded biases.

Our claim is that our rough proxies work better than the current system's measures for three reasons.

First, the measures we propose are relatively apolitical and objective. When a judge works harder, produces quality opinions, is cited by her colleagues, and makes decisions independently of a political party, she is likely to do better in the tournament. Performance based on these measures is not about pleasing one's political masters.

Second, our measures focus on predicting important aspects of judicial behavior that are overlooked by the current system. For example, little or no attempt is made currently to measure objectively the quality of a candidate's writing. We find this odd because writing opinions is a key part of an appellate judge's job.

Third, the system we propose produces less of the harmful game-playing than the current one (where judges may skew their opinions to fit into a political position favored by the president). If the game rewards high-quality work and a player's method of behaving strategically is to work harder and produce better quality work, then gaming is a good thing. Game-playing is bad only when the game creates the wrong incentives by rewarding the wrong outcomes. Consider the current system. It provides incentives for judges to signal their political ideology by voting on certain hot-button issues in a way that pleases their political masters. On one level, such a signal is easy to send. For example, one simply votes a particular way on an abortion or death penalty case. The complication, however, is that judges seeking advancement will have to be careful not to signal their ideology too strongly to the other side because of the danger that it will muster its resources to block them. The game-playing that will go on, therefore, will provide stealth signals to one's political masters, while sending ambiguous or misleading signals to one's opponents.[15] The costs are reduced transparency and added uncertainty. Politicians will push

15. A concrete example of this might be a Republican judge who avoids writing a dissent in an anti-death penalty decision out of a fear that the dissent might be picked up as a signal of strong ideology by the other side. This avoidance has costs. It denies the losing party the dissent that might have helped them in an appeal. It hurts the quality of the majority opinion in that the majority opinion is likely to be better written and argued if there is the threat of a dissent. And, perhaps most important, it denies the public information about the judge's true voting preferences—information that would help predict how the judge would perform if elevated.

candidates with relatively blank records, with the hope that they will turn out to support the ideologies of their sponsors. The result will be uncertainty in the candidate's future performance in terms of both ideology and quality. The tournament we propose has the potential to reduce uncertainty on both scores.

The American Bar Association ("ABA") already conducts an evaluation of potential candidates for the Supreme Court. What justifies a tournament if we already have the ABA's assessment of merit? The easy answer is that the current administration has decided not to use the ABA's evaluations.[16] But there is more to this than a simple choice by the current administration. Oversimplifying, under the ABA process, a group of elite and supposedly nonpartisan lawyers conduct a highly rigorous inquiry into a candidate's background. They evaluate the candidate in terms of three broad criteria: integrity, professional competence, and judicial temperament.[17] As part of the inquiry, the candidate is interviewed and asked to answer a questionnaire. Numerous others who have interacted with the candidate are also interviewed. The ABA group then evaluates the candidate's work product and views, including those reflected in the candidate's articles, other writings, speeches, and legal briefs.[18] From what the ABA reports, the inquiry is exhaustive and thorough. In the past, both presidents and senators have attached significant weight to the rankings that the ABA reports.[19] (It ranks the candidates as either well-qualified, qualified, or not qualified.)[20]

In sum, the ABA may well have a better process to evaluate judicial merit than we do. Our simple measures cannot get directly at what we readily concede are important elements of judicial merit such as integrity and temperament. But the point of our tournament is not to come up with a

16. *See* Laura E. Little, *The ABA's Role in Prescreening Federal Judicial Candidates: Are We Ready to Give Up on the Lawyers?*, 10 WM. & MARY BILL RTS. J. 37, 37 (2001–2002).

17. AMERICAN BAR ASSOCIATION, THE ABA STANDING COMMITTEE ON FEDERAL JUDICIARY: WHAT IT IS AND HOW IT WORKS 4 (1999), *at* http://www.abanet.org/poladv/scfedjud.pdf [hereinafter ABA REPORT].

18. *Id.* at 8–9 (describing the vetting process for Supreme Court candidates in particular).

19. *See* HENRY J. ABRAHAM, JUSTICES, PRESIDENTS, AND SENATORS 23 (revised ed. 1999). *Cf.* ABA REPORT, *supra* note 17, at 8 ("The Committee has been consulted by every President concerning virtually every federal judicial appointment since 1952, and the United States Senate, through the Senate Judiciary Committee, has been provided with the Committee's evaluation of every federal judicial nomination since 1940.").

20. ABA REPORT, *supra* note 17, at 8. For more detail on the role that the ABA's evaluations have played historically, see, for example, ABRAHAM, *supra* note 19, at 23–28; GEORGE WATSON & JOHN A. STOOKEY, SHAPING AMERICA: THE POLITICS OF SUPREME COURT APPOINTMENTS 83–85 (1995).

perfect (or best) measure of judicial merit. It is to flush out ideological motivations. The ABA's system cannot do this because its process is shrouded in secrecy.[21] Its claim that its system of ranking is credible, accurate, and free from bias is entirely wrapped up in its elite institutional status, the reputations of its members, and some notion of professionalism. To the extent that this reputational bond works to guarantee the unbiased quality of the evaluation, that is fine. The reputational bond, however, is not working anymore. The current administration sees the ABA as having a liberal bias, and there is some empirical evidence to back their suspicions.[22] The value of our tournament in comparison to the ABA ranking system is that where the ABA evaluation process and eventual ranking is nontransparent and unverifiable, ours is designed to be precisely the opposite of those things. The ABA's criteria are of little use in distancing the politicians' claims of merit from suspicions about ideology if the ABA's rankings themselves are based on subjective (and somewhat secret) criteria. We do not doubt that our tournament is also subject to claims of bias. The transparent nature of our objective measures, nonetheless, allows outside observers to both critique the tournament and adjust the rankings for such biases.

A. ADOPTING A MERIT-BASED TOURNAMENT

Would the president be willing to adopt merit-based criteria, using objective measures like those we propose, to evaluate the worthiness of a Supreme Court nominee? The president's power to nominate a Justice is constitutionally derived. The Constitution gives the president the discretion to nominate, subject only to the advice and consent of the Senate. But the Constitution says nothing about what standards are to be used. The Appointments Clause provides that the president "shall nominate, and by

---

21. *See* ABA REPORT, *supra* note 17, at 10, 12 (stressing the importance of confidentiality in the vetting process and declining to provide precise evaluative criteria or methodology). This need for transparency would apply to a White House selection process that involved detailed screening. Indeed, the need for transparency would be even more applicable there to the extent that the public is more suspicious of the White House's incentives to screen based on political litmus tests rather than merit. *See* MALTESE, *supra* note 4, at 121–23 (describing the detailed screening process used by the Reagan White House in which candidates were put through four or five hours of screening interviews with multiple White House attorneys, and both their speeches and prior writings were carefully scrutinized).

22. *See* James Lindgren, *Examining the American Bar Association's Ratings of Nominees to the U.S. Courts of Appeals for Political Bias, 1989–2000*, 17 J.L. & POL. 1 (2001). *But see* John R. Lott, Jr., *The American Bar Association, Judicial Ratings, and Political Bias*, 17 J.L. & POL. 41 (2001) (finding only weak evidence of bias); Michael J. Saks & Neil Vidmar, *A Flawed Search for Bias in the American Bar Association's Ratings of Prospective Judicial Nominees: A Critique of the Lindgren Study*, 17 J.L. & POL. 219, 252–54 (2001) (critiquing Lindgren's methodology and conclusions).

and with the Advice and Consent of the Senate, shall appoint . . . Judges of the supreme Court, and all other Officers of the United States."[23] Michael Gerhardt explains that the Framers drafted the Appointments Clause not to select the most qualified Supreme Court Justices, but to prevent nepotism and tyranny.[24] In so drafting, the Framers provided little guidance on the ideal qualifications to look for in a candidate.[25] Since the Constitution's ratification, formal rules and informal norms have developed to fill out the Appointments Clause.[26] Among these norms is the process of selecting or blocking a candidate based on ideological grounds.[27]

---

23. U.S. CONST. art. II, § 2, cl. 2.

24. Gerhardt writes that

> [c]ontrary to the assumption of many of its critics, the Framers did not design this system to ensure the appointments of the best-qualified people to important governmental offices. Rather, the Framers' primary concern in designing the system was to preclude certain kinds of abusive or inappropriate appointments. Some Constitutional Convention delegates were primarily concerned with developing a system that would protect against legislative tyranny, whereas many others were concerned mostly with preventing monarchic despotism.

Michael J. Gerhardt, *Toward a Comprehensive Understanding of the Federal Appointments Process*, 21 HARV. J.L. & PUB. POL'Y 467, 474–75 (1998).

25. As John Baker explains,

> [t]he Constitution does not lay out particular qualifications for justices of the Supreme Court. It does not require, as it does for members of Congress and the President, that they have attained a certain age. The Constitution does not even require that justices be lawyers. It simply provides that the President "shall nominate, and by and with the Advice and Consent of the Senate, shall appoint . . . Judges of the supreme Court . . . ." Just as it cannot add to the qualifications of its own members beyond those provided in the Constitution, presumably Congress also cannot add to the qualifications for either the President or justices of the Supreme Court.

Baker, *supra* note 7, at 178.

But would we have benefited from more guidance? *See* Lee Epstein, Jack Knight & Olga Shvetsova, *Comparing Judicial Selection Systems*, 10 WM. & MARY BILL RTS. J. 7 (2001–2002) (looking at other countries' constitutionally mandated criteria for judicial appointees and comparing their results with the United States in achieving a distinguished and independent bench).

26. *See generally* Michael J. Gerhardt, *Norm Theory and the Future of the Federal Appointments Process*, 50 DUKE L.J. 1687, 1687 (2001) (discussing the "behavioral regularities of presidents and senators regarding appointments that persist in the absence of formal rules and that deviations from which trigger sanctions"). Norms include, for example, the practice of senatorial courtesy. *See* Brannon P. Denning, *The Judicial Confirmation Process and the Blue Slip*, 85 JUDICATURE 218, 220 (2002). For example, the "blue slip" practice enables a senator to facilitate the withdrawal of a nominee, or "blue slip" the nominee, even before the Senate's review of a nominee begins. *See id.*

27. *See* James J. Brudney, *Recalibrating Federal Judicial Independence*, 64 OHIO ST. L.J. 149, 157 (2003) ("Political and ideological background have been relied upon by presidents, senators and their staffs when choosing from within a pool of talented and experienced candidates."); David M. Levitan, *The Effect of the Appointment of a Supreme Court Justice*, 28 U. TOL. L. REV. 37, 69 (1996) (analyzing instances in which the appointment of a single Supreme Court Justice has directly affected the law through reversals of earlier decisions and concluding that "it is right and proper for a President to seek to determine a nominee's values, attitudes, ideas and motivations before appointing the nominees and for the Senate to do so before confirming the nominee"). As noted earlier, scholars hold different views on whether the Senate should actively use ideology as grounds for refusing to confirm a nominee. *See, e.g.*, Eastman, *supra* note 7, at 647 (arguing that the Senate's confirmation power "exists only to prevent the President from selecting a nominee who 'does not possess due qualifications for

In sum, there appears to be little room for an argument that the Constitution requires the president and Senate to use standards in selecting judicial candidates. Indeed, given the open-ended grant of discretion to the president (in nominating) and the Senate (in its advice and consent role), it is likely that the imposition of standards, especially those based on an objective tournament of judges, would require a constitutional amendment—unless, of course, the president or the Senate were voluntarily to adopt standards. The president and the Senate, however, are unlikely ever to agree to relinquish their discretionary, constitutionally derived power.[28]

Nonetheless, the tournament does not have to be mandatory to have an impact. Publicizing the tournament's methodology and its results introduces a standard and gives the public tools to judge both the president's nominee as well as the Senate's opposition or support of the nominee. Rather than assess a candidate against a vacuum, the public will know what the tournament criteria are and understand why the winners won. And the public may rightly ask why the president chose a nominee not among the winners, if this is, in fact, the case. Pressure will then mount for the president (as well as the Senate) to be more explicit in their ideological motivations. Thus, two unknowns, a nominee's merit and the president's ideological basis for choosing one nominee over another, will be revealed. If the public disagrees with the way the nomination process is proceeding, it will pressure elected officials to justify or abandon their positions. To alleviate public pressure, the president and Senate, in turn, may view those who do better in the tournament with more favor.

In order for the tournament to work, broad stakeholder participation—including participation from the media, civil society, and the general public—is required both to monitor judicial performance (that is, to run the

---

office.' Essentially the Senate's confirmation power exists to prevent the President from being swayed by nepotism or mere political opportunism" and does not allow it to impose ideological litmus tests on candidates (quoting JOSEPH STORY, COMMENTARIES ON THE CONSTITUTION OF THE UNITED STATES § 1525, at 377 (1833))). *Cf. Hearing Before the Senate Committee on the Judiciary Subcommittee on Administrative Oversight and the Courts on Should Ideology Matter?: Judicial Nominations 2001*, 50 DRAKE L. REV. 429, 463–73 (2002) (statement of Cass R. Sunstein) (concluding that ideology should matter, and that it is appropriate for the Senate to ensure a balanced and moderate judiciary in the current era of excessive judicial activism).

28. One scenario in which they might agree voluntarily to restrict their power is where the level of conflict in the appointments process rises to the point where neither side is able to achieve anything. In such a situation, it would be in the interests of both parties to agree to take less power so as to achieve a more cooperative solution. *Cf.* David S. Law, Appointing Federal Judges: The President, the Senate, and the Prisoner's Dilemma (Oct. 29, 2003) (unpublished draft, on file with authors) (modeling the dynamic between the President and the Senate in game theoretic terms).

tournament) and put pressure on politicians based on the resulting judicial rankings. The most identifiable stakeholder groups in our merit-based proposal are members of the legal profession, academics who study the courts, and public interest groups. We think that our tournament will also create stakeholders in the broader general public for the following reason: most everyone is interested in federal government shenanigans and everyone relates to merit-based promotion. The tournament we propose is relatively transparent, meaning that the methodology is easily understood and the results are readily transmittable, resulting in wide availability. Although the public may have difficulty understanding constitutional interpretation, they do understand that, in an ideal world, hard work is a sound basis for a promotion. To the extent that the tournament's objective criteria are in tune with the public's intuitive sense of merit, the criteria will form a focal point for judging candidates for the High Court.

Would an ongoing tournament be sustainable over time? One of the reasons that hot-button issues have become so important may be the public's apathy. This is not necessarily a bad thing. People have much information to process, and processing information is costly in time and money, so they select carefully what issues interest them. The current process of Supreme Court nominations is, as we have said before, shrouded in secrecy, and therefore it is very costly for Joe and Jane Public to process an opinion about a particular judge. Instead, they default to how the judge votes on a few issues that interest them including, for example, gun rights and affirmative action. In addition, these also tend to be the issues in which interest groups are most interested. The result is that these are the issues that become most important to the president and Senate in the nomination process. We propose that by introducing an objective, merit-based tournament, the public is less likely to default to hot-button issues because the cost to obtain information about a nominee's merit is much lower. Much like the *U.S. News & World Report* rankings of law schools and other professional schools have become influential, the very ease of transmitting the tournament's objective criteria will give the tournament's results wide influence among the public.[29]

29. Indeed, the opposite problem is possible. The ease of transmission and understandability of the tournament outcome may result in the public giving too much weight to the objective criteria. While such a problem is theoretically possible, politicians may respond with greater effort at explaining the true "merit" behind their candidates (who fail to score highly on the objective criteria).

III. CONSTRUCTING THE TOURNAMENT

We limit our sample to federal circuit court judges other than those on the Federal Circuit.[30] While in theory, and recently in practice, the president may select a Supreme Court nominee from the pool of state court judges and even from among nonjudges, the norm in recent years has been to select from among the sitting federal circuit court judges from the twelve circuits of general jurisdiction. For purposes of this Article we take that norm as given. While including other types of judges, such as those on the Federal Circuit, is possible, a tournament including such judges would have to adjust for the differences between the number and types of opinions facing these other judges. Finally, we run the tournament with judges from both the major political parties—the assumption being, however, that the sitting president will focus on the top performers from his own party.

As our initial sample, we select only those federal circuit judges still active as of June 2003. Judges who retire or take senior status as of June 2003 are excluded from the tournament. Presumably, the choice of retiring or taking senior status suggests either a diminished capacity for, or interest in, judging. Additionally, we include only those judges who were appointed prior to January 1, 1998. In effect, then, the tournament posits a six year apprenticeship period (given that we calculate data up to May 31, 2003) before one can be considered for promotion to the Supreme Court. The

30. The two circuits that have the most specialized and, therefore, noncomparable dockets are the D.C. Circuit and the Federal Circuit. In excluding the Federal Circuit and including the D.C. Circuit, we made a judgment call that at least a meaningful portion of the D.C. Circuit's docket was comparable to the other circuits. This comparability factor is most important with respect to citation numbers. So, for example, because of the large number of (often burdensome) administrative law cases that the D.C. Circuit sees compared to other circuits, it may not be meaningful to compare the total citations of a D.C. Circuit judge's published opinions to those of a judge on the First or Seventh Circuits. That said, a comparison of the respective numbers of citations to their top twenty opinions may be more meaningful. Focusing on only a judge's top twenty opinions reduces (if not eliminates) the advantage of judges that write more opinions. Or, if that is found wanting, one could do a comparison of the relative citation numbers in the law reviews. While circuits outside of the D.C. Circuit facing a relative dearth of administrative law cases may cite D.C. Circuit cases less, we suspect that the same does not hold true for the law reviews (where administrative law is a frequent topic of scholarship).

For us, even if our initial measures are flawed, the key is to get a discussion going. So, hypothetically, if we were to find that Judge Harry Edwards ranks thirtieth overall, but that he scores high in productivity and independence and low in citations, a discussion could occur as to whether citation scores are an irrelevant measure for him because of the D.C. Circuit's unusual docket. It might, for example, be argued that Edwards's relative ranking on our invocation scale or his ranking based on citations in law journals should be used. Regardless of how such a discussion goes, we contend that starting from an objective set of criteria will lead to a more informed discussion than one initiated in a vacuum (and one more transparent with respect to ideology). On the historical evolution of the D.C. Circuit's docket, see generally CHRISTOPHER P. BANKS, JUDICIAL POLITICS IN THE D.C. CIRCUIT COURT (1999).

apprenticeship period enables the generation of data on these judges. The tournament focuses on published opinions written during the 1998 to 2000 time period for all the judges. We obtain the published opinions for each judge from Westlaw and Lexis.

Establishing one common time period eliminates the need to control for differences across time as well as differences in the length of service among judges. Judge X may have more opinions and more corresponding citations simply because Judge X has served on the bench longer than Judge Y. Examining only the opinions written by each judge in the tournament over the 1998 to 2000 time period puts the judges on a relatively even playing field. Each judge has the same amount of time to generate opinions. While some judges may write more opinions and receive more citations for their opinions because of age or experience, this is precisely what we are looking for in the tournament—differences attributable to internal differences across judges rather than external differences (such as different time periods). It is possible, but not necessary, that older and more experienced judges will be able to write more opinions and generate more citations than their younger and more inexperienced counterparts. That is not a problem because, if the older and more experienced judges are doing better work, they are the ones who should be promoted. Their younger counterparts will have their chance in later tournaments.

Given these restrictions, our sample consists of ninety-eight federal circuit court judges. The judges are distributed across the circuits as follows:

TABLE 1. Distribution of judges in the tournament by circuit

Circuit	Number of Judges
1st	4
2d	5
3d	7
4th	8
5th	13
6th	8
7th	10
8th	5
9th	12
10th	8

Circuit	Number of Judges
11th	10
D.C.	8
Total	**98**

Judges decide cases and write opinions explaining their decisions. To determine relative performance levels for purposes of the tournament, we need a set of measures that get at the quality and quantity of decisionmaking and opinion-writing. Our measures fall roughly into three categories: those evaluating (a) productivity in providing published statements of reasons ("productivity"), (b) quality of opinion-writing ("quality"), and (c) independence from the views of one's colleagues and political sponsors ("independence").

A. PRODUCTIVITY

The number of cases a circuit judge hears is largely a function of the circuit on which the judge sits. There are significant differences in the caseloads across the circuits, but the one commonality is that the burdens are overwhelming.[31] They are so overwhelming that almost no judge can hope to provide a publication-worthy statement of reasons in every case that comes before the judge. Some judges, however, provide published statements of reasons in more cases than others.

Presumably, it takes greater effort and skill to write more published opinions.[32] Measuring effort exerted by a judge is an important element of the tournament because comparing the past effort levels of the various judges (a) helps predict future effort levels (we want Justices who will exert high levels of effort) and (b) helps determine who among the lower court judges should be rewarded for their efforts (so as to show lower court judges that their high effort levels are valued and will be rewarded). The

31. The explosions in the caseloads of the federal courts and the various strategies used to tackle the overload have been documented by a number of commentators. *See, e.g.*, Jeffrey O. Cooper & Douglas A. Berman, *Passive Virtues and Casual Vices in the Federal Courts of Appeals*, 66 BROOK. L. REV. 685 (2000–2001).

32. *See* Alex Kozinski & Stephen Reinhardt, *Please Don't Cite This!*, CAL. LAW., June 2000, at 43 (explaining the greater effort that goes into writing published opinions compared to the effort that goes into unpublished memorandum dispositions). *See also* William Glaberson, *Caseload Forcing Two-Level System for U.S. Appeals*, N.Y. TIMES, Mar. 14, 1999, at A1 (quoting Judge Posner for the proposition that most judges are not as careful with unpublished dispositions as they are with published decisions).

measure of effort we use is the number of published opinions from 1998 to 2000.

Judges also write unpublished opinions. Indeed, the majority of opinions these days are unpublished.[33] Unpublished opinions represent opinions judges affirmatively do not want to be used by others as precedents. The implication, and our assumption, is that they often involve minimal effort (and a lower quality of reasoning). Given our desire to measure the willingness to exert a high level of effort, we focus on published opinions. Judges may also demonstrate their productivity in other ways. Some judges may engage in an active public speaking schedule. Others may write academic articles or teach the occasional law school class. While such pursuits may be valuable, we focus solely on published opinions for three reasons. First, the number of published opinions may well be correlated with greater numbers of law review articles and other forms of communication to the public. Judge Posner, a prolific author of academic articles and books, also consistently publishes the largest number of judicial opinions per year.[34] Second, one of the most important functions for a judge is generating opinions to serve as precedents for others (that is, reducing the amount of uncertainty in the law). Giving a judge credit for doing other things will diminish the incentive to spend time on opinion writing.

Table 2 reports the total number of published opinions (consisting of majority, concurring, and dissenting opinions) for the ten judges with the greatest number of opinions. We also report the total number of majority opinions published by each judge in Table 2. The Appendix provides the ranking for the entire set of tournament judges (reported in Appendix Table B).

---

33. The practice of issuing unpublished opinions in the courts of appeals began in the 1970s and today accounts for more than seventy-five percent of all decisions. *See* Stephen L. Wasby, *Unpublished Decisions in the Federal Courts of Appeals: Making the Decision to Publish*, 3 J. APP. PRAC. & PROCESS 325, 325 (2001).

34. *Cf.* Fred R. Shapiro, *The Most-Cited Legal Scholars*, 29 J. LEGAL STUD. 409, 424 tbl. 6 (2000) (reporting that Judge Posner is among the most cited of the jurisprudential giants, including those such as Justice Oliver Wendell Holmes and Roscoe Pound). For research reporting the average number of opinions signed by Judge Posner per year, see Landes et al., *supra* note 13, at 278 n.20.

TABLE 2. Published opinions, 1998–2000 (for the ten judges with the highest number of published opinions)

	(A) Total Number of Published Opinions	(B) Total Number of Published Majority Opinions	(C) Circuit
Richard Posner	**269**	**254**	7th
Frank Easterbrook	**233**	**213**	7th
Joel Flaum	202	192	7th
Diane Wood	194	173	7th
Kenneth Ripple	182	151	7th
Michael Kanne	177	176	7th
Morris Arnold	175	152	8th
John Coffey	168	162	7th
James Loken	167	147	8th
Roger Wollman	158	154	8th

The highest two numbers in each category in bold type. Summary statistics for (A) (n=98): Mean = 98.1; Median = 85.5; Standard Deviation = 42.8; Kurtosis = 2.501; Skewness = 1.418.

Summary statistics for (B) (n=98): Mean = 83.6; Median = 74.0; Standard Deviation = 41.5; Kurtosis = 2.918; Skewness = 1.576.

Chi-squared test of the null hypothesis that the distribution of circuits is identical between the top judges and the bottom judges for the number of published opinions (A): $\chi^2 = 34.697$ (11 d.f.) ($p \leq 0.0003$). Top judges are defined as those who are in the top fifty percent of judges in the entire sample (n=98) based on the number of published opinions (majority, concurrences, and dissents) (A). Bottom judges are defined as those who are in the bottom fifty percent of judges based on the number of published opinions (majority, concurrences, and dissents) (A).

The mean number of total published opinions for all the judges in the sample is 98.1 opinions (an average of 32.7 opinions per year). Judge Posner and Judge Frank Easterbrook, the two publishers of the largest numbers of opinions, are each more than three standard deviations above the mean. The next two highest, Judge Joel Flaum and Judge Diane Wood, are both more than two standard deviations above the mean. The mean number of published majority opinions is 83.6. As with the number of total published opinions, both Judge Posner and Judge Easterbrook have the greatest number of published majority opinions (and are again each over three standard deviations above the mean).

A caveat is that these publication numbers are likely driven not only by individual effort, but by additional factors such as a court's culture (the ethic on some courts may be to publish more opinions) and the court's

caseload. [35] Circuits may exhibit different norms that govern when to publish an opinion. Additionally, some circuits may have a norm of deciding more cases in less time or they may face a greater or more complex caseload. Judges from a low productivity circuit may switch into a higher mode of productivity if placed in a different circuit. All the judges in Table 2 are either from the Seventh or Eighth Circuits. In particular, the Seventh Circuit, with seven of the top eight opinion publishing judges in the sample, may represent the results of a high-publishing norm circuit. The chi-squared test rejects the null hypothesis that the distribution of circuits for the top judges and the bottom judges (distinguished based on the number of published opinions) is identical ($\chi^2 = 34.697$ (11 d.f.) ($p \leq 0.0003$)). Statistically significant variation, therefore, exists across the circuits in terms of number of total published opinions.

The fact that a norm to publish more opinions exists in some circuits does not detract from the accomplishment of a judge who does, in fact, produce a large number of publishable opinions. A high-publication norm does not, in itself, make it any easier for a judge to research and write any particular opinion. On the other hand, judges in a circuit with a high-publication norm may compensate through greater reliance on the legal reasoning and research of law clerks, interns, and staff attorneys.[36]

To control for the possible influence of circuit-based norms on publication rates, we determine the mean number of opinions published for the judges of each circuit. The mean number of total opinions for judges of the Seventh Circuit is 185.2 (the highest among all the circuit means). In

35. A number of scholars have observed the importance of circuit norms in determining aspects of judicial publication practices. *See, e.g.*, Deborah Jones Merritt & James J. Brudney, *Stalking Secret Law: What Predicts Publication in the United States Courts of Appeals*, 54 VAND. L. REV. 71, 76–79, 85–90 (2001); Ahmed E. Taha, *Publish or Paris: Evidence of How Judges Allocate Their Time*, 6 AM. L. & ECON. REV. 1 (2004). The Seventh Circuit, for example, appears to have a culture of publishing a high fraction of its opinions and the Third Circuit, by contrast, appears to have the opposite culture. *See* Mitu Gulati & C.M.A. McCauliff, *On* Not *Making Law*, LAW & CONTEMP. PROBS, Summer 1998, at 157, 220 tbl. VIII. Along these lines, some readers may notice that the top scorers in both the number of published opinions and the number of published pages are from the Seventh Circuit (Judges Diane Wood, Posner, and Easterbrook). But one should not be too quick to discount their high scores as a function of the circuit's norms. This is because the norm, in turn, is a function of the inclinations of the individuals on the court. For example, it is likely that the Seventh Circuit's norm is due, in part, to the influence of Judge Posner (and then Judge Easterbrook and perhaps now Judge Wood). *See* Mitu Gulati & Veronica Sanchez, *Giants in a World of Pygmies? Testing the Superstar Hypothesis with Judicial Opinions in Casebooks*, 87 IOWA L. REV. 1141, 1180–81 (2002). From this vantage point, it can be argued that Judges Posner, Easterbrook, and Wood should receive more credit, and not less, for being on a circuit that has a norm of high production.

36. The ability to utilize effectively law clerks and staff attorneys to produce a greater quantity of published opinions is not necessarily a bad thing. To the extent that such management skills are important on the Supreme Court, this might be a positive.

contrast, the mean number of total opinions for the Third Circuit is 60.1. For each circuit other than the Seventh Circuit, we calculate the difference between the other circuit's mean and the Seventh Circuit's. We then add this mean difference to the number of opinions for each judge in the other circuit.

For example, the Third Circuit total opinions mean is 125.1 less than the Seventh Circuit total opinions mean. Judge Samuel Alito, a judge on the Third Circuit, wrote seventy opinions from 1998 to 2000. We adjusted Judge Alito's total opinion count upward to 195.1. After the adjustment, all the circuits have the same mean number of total opinions written. Any differences among judges will, therefore, be determined solely by each judge's standing relative to the other judges within her own circuit. For example, if Judge Morris Arnold is relatively far above his peers in the Eighth Circuit (compared with how much higher the most productive judges in other circuits are above their intracircuit peers), he will still receive a high adjusted ranking for productivity. Table 3 reports the top ten judges based on the number of total published opinions adjusted for intercircuit differences. (Appendix Table C reports the ranking of all judges in the sample based on total published opinions adjusted for intercircuit differences.)

TABLE 3. Published opinions, 1998–2000 (for the ten judges with the highest number of published opinions), adjusted for intercircuit differences

	(A) Total Number of Published Opinions (adjusted for circuit variation)	**Z-Score of (A)**	**(B) Circuit**
Richard Posner	269	3.60**	7th
Stephen Reinhardt	237	2.23**	9th
Diarmuid O'Scannlain	234	2.10**	9th
Frank Easterbrook	233	2.05**	7th
Karen Nelson Moore	231	1.94	6th
Ronald Lee Gilman	225	1.69	6th
Gerald Bard Tjoflat	224	1.66	11th
Jerry Smith	223	1.63	5th

	(A) Total Number of Published Opinions (adjusted for circuit variation)	Z-Score of (A)	(B) Circuit
Paul Niemeyer	221	1.54	4th
Dolores Sloviter	217	1.37	3d

** Indicates a z-score of 1.96 or higher (representing a two-sided probability of <5% for a normal distribution). The number of published opinions for each judge is adjusted so that the mean number of total opinions for each circuit is identical and equal to 185.2 (the unadjusted mean number of total opinions for the Seventh Circuit).

Summary statistics for (A) (n=98): Mean = 185.2; Median = 182.05; Standard Deviation = 23.296; Kurtosis = 1.263; Skewness = 0.543.

Chi-squared test of the null hypothesis that the distribution of circuits is identical between the top judges and the bottom judges for the number of published opinions adjusted for intercircuit differences (A): $\chi^2 = 5.253$ (11 d.f.) ($p \leq 0.918$). Top judges are defined as those who are in the top fifty percent of judges in the entire sample (n=98) based on the number of published opinions (majorities, concurrences, and dissents) adjusted for intercircuit differences (A). Bottom judges are defined as those who are in the bottom fifty percent of judges based on the number of published opinions (majorities, concurrences, and dissents) adjusted for intercircuit differences (A).

The adjusted rankings capture solely variations within a circuit.[37] Even with the adjustment for intercircuit differences in means, note that Judge Posner is still the highest-scoring judge in terms of productivity, and the difference in Judge Posner's production from the mean judge is statistically significant. After correcting for the greater mean number of opinions published in the Seventh Circuit, Judge Posner's relative standing among other Seventh Circuit judges is high enough to place him ahead of high-scoring judges in other circuits. Judge Easterbrook, however, drops to fourth place behind Judge Stephen Reinhardt and Judge Diarmuid O'Scannlain of the Ninth Circuit (all of whom are also significantly above the mean judge). Put another way, Judge Reinhardt's relative ranking among Ninth Circuit judges is higher than Judge Easterbrook's relative ranking among Seventh Circuit judges.

While productivity is a key factor to consider in selecting a Supreme Court Justice, other factors exist. Starting with easily measurable criteria, nonetheless, helps focus on exactly what those other justifications are.

37. The chi-squared test cannot reject the null hypothesis that the distribution of circuits in the top (representing the top fifty percent of judges based on the number of total published opinions and correcting for intercircuit differences) and bottom groups of judges (the bottom fifty percent) is identical ($\chi^2 = 5.253$ (11 d.f.) ($p \leq 0.918$)).

B. MEASURING OPINION QUALITY

A central component of a Justice's role on the Supreme Court is writing opinions. These opinions form the primary basis for the public's understanding of the current state of the law. Predicting the quality of opinions that a judge is likely to write if promoted, therefore, should presumably be a key part of any promotion decision. Writing opinions is also a key element of a circuit court judge's job. Hence, encouraging the writing of better opinions by rewarding those who produce high-quality opinions is important. The question is how to rank judges on the quality of their opinions.

The opinions that any judge writes are cited by other judges to help in explaining the other judges' subsequent decisions. The opinions are also used by scholars and commentators to explain and analyze the law for clients, other scholars, and other lawyers. Thus, there are at least two sets of customers who use opinions. Some opinions will help explain the law better than others and customers will presumably use these opinions more. Examining customer use provides a market test of the quality of a judge's opinions.[38]

---

38. For a discussion of citation measures to rank judicial influence and the various problems using such measures, see Landes et al., *supra* note 13, at 271–76. Additional discussions of the limitations of using citation counts to measure judicial influence are found in Virgil L.P. Blake, *Citation Studies—the Missing Background*, 12 CARDOZO L. REV. 1961 (1991) (book review), and Keith Ann Stiverson & Lynn Wishart, *Citation Studies—Measuring Rods of Judicial Reputation?*, 12 CARDOZO L. REV. 1969 (1991) (book review). On citation analysis more generally, see Richard Posner, *An Economic Analysis of the Use of Citations in the Law*, 2 AM. L. & ECON. REV. 381 (2000). *See also* Symposium, *Interpreting Legal Citations*, 29 J. LEGAL STUD. 317 (2000); Symposium, *Trends in Legal Citations and Scholarship*, 71 CHI.-KENT L. REV. 743 (1996).

The use of citations as a proxy for research quality has been discussed by scholars in a number of other areas. For studies that either suggest or assume a relationship between citation counts and quality, see Luis R. Gomez-Mejia & David B. Balkin, *Determinants of Faculty Pay: An Agency Theory Perspective*, 35 ACAD. MGT. J. 921 (1992) (devising a proxy for research productivity in management science with a qualitative-quantitative measure of citations and journal quality to assess determinants of faculty pay); Michael E. Gordon & Julia E. Purvis, *Journal Publication Records as a Measure of Research Performance in Industrial Relations*, 45 INDUS. & LAB. REL. REV. 194, 195 (1991) (using citations as an objective measure of "research excellence"); J. Scott Long, *Productivity and Academic Position in the Scientific Career*, 43 AM. SOC. REV. 889, 892 (1978) (using number of publications and citations as a proxy for faculty productivity in the field of biochemistry); Kee H. Chung, Raymond A.K. Cox & John B. Mitchell, *Quantity vs. Quality: Impact on Scholarly Contribution?* (Central Michigan University, Working Paper No. 10/17/03, 2003), *at* http://ssrn.com/abstract=459742 (examining the relationship between the number of papers published by an individual, what institution published them, and the number of citations the individual receives); Scott Smart & Joel Waldfogel, *A Citation-Based Test for Discrimination at Economics and Finance Journals* (NBER Working Paper No. 5460, 1996), *at* http://ssrn.com/abstract=225512 (using citations as a measure of article quality, and then asking whether papers by members of certain groups receive systematically different numbers of citations). *But see* Waldo C. Klein & Martin Bloom, *Studies of Scholarly Productivity in Social Work Using Citation*

For each judge, we collected data from Shepard's (on the Lexis database) on the number of times the judge's opinions published from 1998 to 2000 were cited in other judicial opinions and secondary sources. Focusing on the opinions published in the same time period puts the judges on a relatively level playing field. In theory, each judge has the ability to write opinions of the same quality and receive the same number of citations to such opinions.

Even with the same time period to generate opinions, we may predict that certain judges will receive a greater number of citations than other judges. Judges who are better liked, who are on more respected circuits, who have been on the bench longer, or who have the favor of the current Supreme Court may receive more citations.[39] Judge Easterbrook, for example, has a history of being well-cited in many law periodicals and judicial opinions. The large number of citations Judge Easterbrook has received in the past is due, at least in part, to the value of his analyses and the clarity of his presentation. Judge Easterbrook's past then, makes it more likely that his opinions during the 1998 to 2000 time period will be cited on reputational grounds alone. Of course, if his opinions are of a terrible quality, that reputational presumption will probably be nullified. Our methodology does not control for such pretournament inherent differences. On the contrary, it is precisely this reputation for quality analysis (and the underlying ability behind such a reputation) that we hope to find in a Supreme Court nominee.

Citations come in a variety of forms: citations in a judge's own jurisdiction (where opinions have binding authority), citations by outside courts (including the Supreme Court, other circuit courts, district courts in other circuits, and state courts), and self-citations (in which a judge cites her own opinion). Some measures are more indicative of opinion quality than others. For example, citations by courts in other jurisdictions are more indicative of a high-quality opinion than are citations from one's own jurisdiction. This is because courts in the same jurisdiction often have to

---

Analysis, 28 J. SOC. WORK ED. 291 (1992) (suggesting differing citation norms in social work among academic positions at national and local levels); Robert West & Ann McIlwaine, *What do Citation Counts Count for in the Field of Addiction? An Empirical Evaluation of Citation Counts and Their Link with Peer Ratings of Quality*, 97 ADDICTION 501 (2002) (evaluating the value of citation counts as a proxy for quality in the addiction field and finding no correlation between the number of citations and independent expert ratings of article quality, but finding a correlation between citation counts and geographic region).

39. A method of correcting for such biases is to use regression analysis that allows one to separate out the effects of the different factors such as years on the bench, circuit affiliation, and the sort. For a study that conducts such an analysis, see Landes et al., *supra* note 13.

cite intracircuit opinions as binding precedent. Citations to outside circuits (termed "outside-circuit citations"), however, occur only if citing the opinion serves a purpose in making an argument.[40]

Table 4 reports summary citation results. (Appendix Table D reports for the entire sample of judges.) Focus on column (A) for the total numbers of outside court citations, defined to include citations from other circuit courts (outside of a judge's home circuit), state court citations, and U.S. Supreme Court citations. Citation data are gathered from Shepard's on Lexis and measured as of May 31, 2003. Judge Posner and Judge Easterbrook dominate (and are significantly different from the mean judge). The mean judge in the sample received 417.3 outside-circuit citations. Judge Posner and Judge Easterbrook each are over four standard deviations above the sample mean.

TABLE 4. Citations to opinions published, 1998–2000 (for the ten judges with the highest total number of outside citations)

	(A) Total Outside-Circuit Citations	Z-Score of Normalized (A)	(B) Supreme Court Citations	Z-Score of Normalized (B)	(C) Law Review and Periodical	Z-Score of Normalized (C)	(D) Self-Citations	Z-Score of Normalized (D)	(E) Circuit
Richard Posner	1406	2.61**	16	2.31**	1033	2.41**	392	2.35**	7
Frank Easterbrook	1340	2.52**	14	2.11**	790	1.83	257	1.95*	7
Sandra Lynch	1023	1.99**	5	0.62	684	1.52	178	1.60	1
Bruce Selya	949	1.85	3	-0.04	727	1.65	364	2.28**	1
Paul Kelly	799	1.51	0	-2.29**	388	0.30	103	1.07	10
Michael Kanne	768	1.44	4	0.32	512	0.90	181	1.61	7
Joel Flaum	743	1.37	3	-0.04	613	1.29	126	1.27	7
Kenneth Ripple	730	1.34	4	0.32	545	1.03	168	1.54	7

40. *See id.* at 272–73 (noting that "citations to an opinion from within a circuit may reflect either the opinion's precedential or persuasive effect, while citations to an opinion from another circuit will reflect its persuasive effect alone").

	(A) Total Outside-Circuit Citations	Z-Score of Normalized (A)	(B) Supreme Court Citations	Z-Score of Normalized (B)	(C) Law Review and Periodical	Z-Score of Normalized (C)	(D) Self-Citations	Z-Score of Normalized (D)	(E) Circuit
Diane Wood	678	1.20	3	-0.04	513	0.90	127	1.27	7
J. Harvie Wilkinson	662	1.15	4	0.32	648	1.41	23	-0.36	4

** Indicates a Z-score of 1.96 or higher (representing a two-sided probability of <5% for a normal distribution). Outside-circuit citations are measured as of May 31, 2003. Normalized (A) is equal to LN(Total Outside-Circuit Citations). Normalized (B) is equal to LN(1+Supreme Court Citations). Normalized (C) is equal to LN(Law Review and Periodical Citations). Normalized (D) is equal to LN(Self Citations).

Summary statistics for (A) (n=98): Mean = 417.3; Median = 383.0; Standard Deviation = 229.5; Kurtosis = 5.028; Skewness = 1.795.

Summary statistics for normalized (A) (n=98): Mean = 5.903; Median = 5.948; Standard Deviation = 0.515; Kurtosis = 0.025; Skewness = -0.020.

Summary statistics for (B) (n=98): Mean = 3.837; Median = 4.000; Standard Deviation = 2.757; Kurtosis = 4.583; Skewness = 1.547.

Summary statistics for normalized (B) (n=98): Mean = 1.410; Median = 1.609; Standard Deviation = 0.616; Kurtosis = 0.438; Skewness = -0.657.

Summary statistics for (C) (n=98): Mean = 374.2; Median = 375.0; Standard Deviation = 172.0; Kurtosis = 1.408; Skewness = 0.992.

Summary statistics for normalized (C) (n=98): Mean = 5.822; Median = 5.927; Standard Deviation = 0.464; Kurtosis = -0.497; Skewness = -0.148.

Summary statistics for (D) (n=98): Mean = 56.51; Median = 30.50; Standard Deviation = 69.05; Kurtosis = 9.287; Skewness = 2.807.

Summary statistics for normalized (D) (n=98): Mean = 3.508; Median = 3.418; Standard Deviation = 1.049; Kurtosis = 0.509; Skewness = -0.149.

Chi-squared test of the null hypothesis that the distribution of circuits is identical between the top judges and the bottom judges for the number of outside-circuit citations (A): $\chi^2 = 31.553$ (11 d.f.) ($p \leq 0.001$). Top judges are defined as those who are in the top fifty percent of judges in the entire sample (n=98) based on the number of outside-circuit citations (A). Bottom judges are defined as those who are in the bottom fifty percent of judges based on the number of outside-circuit citations (A).

As alternative measures of the quality of opinions, we focus on the number of citations by the U.S. Supreme Court as well as law reviews and periodicals. Columns (B) and (C) of Table 4 report these criteria for the judges with the highest number of outside-circuit citations (collected from Shepard's). Judges Posner and Easterbrook again receive the highest two rankings for both U.S. Supreme Court citations and law review and periodical citations.

William Landes, Lawrence Lessig, and Michael Solimine suggest that judges who write more of their opinions, without the help of law clerks, typically engage in more self-citations (due to their greater familiarity with their self-authored opinions).[41] Judges who contribute a large amount to their opinions consistently over several years typically minimize the input of their law clerks. In other words, judges who rely heavily on their clerks will produce opinions of a more varying quality over the years due to the influences of their clerks. Even a high-quality judge will have difficulty maintaining a high citation count consistently over the years when relying heavily on clerks. One might expect that judges who have a larger number of self-citations (indicating more authoring of opinions) will also score higher on the total citation count. Column (D) of Table 4 reports the self-citation numbers for the top ten total outside citation-receiving judges. The mean number of self-citations for the sample is 56.5 with a standard deviation of 69.1. Note that almost all the judges in Table 4 are well above the mean (except for Judge J. Harvie Wilkinson). Judge Posner and Judge Easterbrook are both at least three standard deviations above the mean (and the difference is statistically significant).

As discussed above, the number of opinions a judge publishes is a matter of choice. Judges have substantial discretion in choosing whether to provide a published explanation of reasons or decide a case with minimal explanation (and sometimes not even that).[42] Other things equal, a judge that authors a larger number of opinions will have a larger number of citations. Where a circuit has a norm of not publishing many opinions, judges may be at a disadvantage not only in the count of total opinions published but also in the number of citations. As evidence of the impact of variation in publication norms on citation counts, the chi-squared test rejects the null hypothesis that the distribution of circuits for top and bottom judges (divided based on the number of outside-circuit citations) are identical ($\chi^2 = 31.553$ (11 d.f.) ($p \leq 0.001$)).[43]

---

41. *See id.* at 274.

42. *See* Gulati & McCauliff, *supra* note 35, at 159, 164.

43. Consistent with our finding of circuit-specific differences in outside-circuit citations, the First and Seventh Circuits do particularly well in the citation study conducted by Landes, Lessig, and Solimine. *See* Landes et al., *supra* note 13, at 302–03.

TABLE 5. Outside-circuit citations to opinions controlling for total number of opinions published, 1998–2000 (for the ten judges with the highest number of outside-circuit citations to each judge's top twenty opinions)

	(A) Outside-Circuit Citations to Top Twenty Opinions	Z-Score of Normalized (A)	(B) Average Outside-Circuit Citations per Majority Opinion	Z-Score of Normalized (B)	(C) Circuit
Sandra Lynch	734	2.56**	9.03	1.73	1st
Frank Easterbrook	667	2.33**	6.25	0.73	7th
Paul Kelly	654	2.28**	9.85	1.97**	10th
Richard Posner	570	1.95	5.49	0.37	7th
Bruce Selya	516	1.71	6.50	0.83	1st
Anthony Scirica	496	1.61	14.50	3.04**	3d
Frank Hull	460	1.43	10.90	2.25**	11th
Karen Williams	455	1.40	11.02	2.28**	4th
Edward Earl Carnes	444	1.34	8.92	1.70	11th
J. Harvie Wilkinson	425	1.24	7.64	1.28	4th

** Indicates a Z-score of 1.96 or higher (representing a two-sided probability of <5% for a normal distribution). Outside-circuit citations measured as of May 31, 2003. Normalized (A) is equal to LN(Outside-Circuit Citations to Judge's Top Twenty Opinions). Normalized B is equal to LN(Average Outside-Circuit Citations per Majority Opinion).

Summary statistics for (A) (n=98): Mean = 277.9; Median = 256.5; Standard Deviation = 121.2; Kurtosis = 2.608; Skewness = 1.382.

Summary statistics for normalized (A) (n=98): Mean = 5.543; Median = 5.547; Standard Deviation = 0.412; Kurtosis = 0.098; Skewness = 0.068.

Summary statistics for (B) (n=98): Mean = 5.137; Median = 4.861; Standard Deviation = 2.030; Kurtosis = 4.577; Skewness = 1.638.

Summary statistics for normalized (B) (n=98): Mean = 1.569; Median = 1.581; Standard Deviation = 0.364; Kurtosis = 0.291; Skewness = 0.216.

Chi-squared test of the null hypothesis that the distribution of circuits is identical between the top judges and the bottom judges for the number of outside-circuit citations to the top twenty citation-receiving cases (A): $\chi^2 = 15.466$ (11 d.f.) ($p \leq 0.169$). Top judges are defined as those who are in the top fifty percent of judges in the entire sample (n=98) based on the number of outside-circuit citations to the top twenty citation-receiving cases (A). Bottom judges are defined as those who are in the bottom fifty percent of judges based on the number of outside-circuit citations to the top twenty citation-receiving cases (A).

One possible correction is to look at the average citations per majority opinion. These numbers are reported in column (B) of Table 5. The problem with this measure is that it gives less credit to a judge who writes several outstanding opinions in addition to many smaller opinions than to a judge who writes the same number of outstanding opinions, but no additional smaller opinions. To get a sense of the number of citations per opinion while both (a) controlling for the total number of opinions and (b) giving due weight to judges who write outstanding opinions and then add on smaller opinions, we look at the total outside-circuit citations to each judge's top twenty citation-receiving opinions. So, regardless of the fact that Judge Posner wrote far more opinions than Judge Sandra Lynch during the 1998–2000 time period, we report only the outside-circuit citations to each judge's top twenty citation-receiving opinions. Column (A) of Table 5 reports this number for the ten highest-scoring judges. (Appendix Table E reports on the entire sample of judges.)

Looking at the same number of opinions for each judge controls for the possibility of a judge using a large number of opinions to generate a high citation count. In addition, the influence of intercircuit differences in opinion publishing norms is minimized (to the extent that the same number of opinions, twenty, is used for each judge regardless of circuit). In support, the chi-squared test cannot reject the null hypothesis that the distributions of circuits are identical for the top judges (based on the number of outside citations to the top twenty opinions) and the bottom judges ($\chi^2 = 15.466$ (11 d.f.) ($p \leq 0.169$)). In addition, given that Justices on the Supreme Court write far fewer opinions than those on the lower courts (and that we are looking at these numbers for a promotion to the High Court), one may argue that it is the judges' performance on their best opinions that we should look at and not just their average performances on a full set of opinions.

Here the two highest-scoring judges for the top twenty opinion citation count are Judge Lynch of the First Circuit and Judge Easterbrook. Judge Posner, while still in the top five, is no longer the highest-scoring judge. In terms of approach, one might infer that Judge Lynch focuses more on crafting high citation opinions than Judge Posner. Judge Posner's individual opinions are not as heavily cited as Judge Lynch's opinions. Judge Posner's sheer productivity, however, results in more overall citations than Judge Lynch.

The use of citations as a proxy for quality will strike some as outrageous. Many believe that it is outrageousness or originality, as

opposed to quality, that gets an opinion cited more. We are skeptical that this is the case, but the question is a fair one. A finer tuned measure of citations could look to reward only those citations that were for authority, as opposed to citations in which the judge is attempting to distance himself or herself from someone else's outrageous statements. Along these lines, a finer tuned measure could also eliminate negative citations, that is, citations of disagreement with a position.[44] Finally, to the extent that it appears law journal articles are more likely to cite outrageous opinions, these numbers can be discounted (and are not included in the citation measures in this Article aside from column (C) in Table 4).

To test whether outrageousness (in a negative way) is driving the high citation counts for the top judges, we examined the number of negative citations (as identified by Shepard's) for each judge's twenty opinions that received the most outside-circuit citations. For comparison, we examined the number of negative citations to the top twenty citation-receiving opinions for the median judge as well as the immediate five judges above and below the median (for a total of eleven comparison judges). Table 6 reports the negative citation counts.

---

44. One of our colleagues who is against using citation measures argues that no one will ever construct finer tuned measures because they are too difficult to do. We disagree. If people find the idea of using such data attractive (and we concede that is a big "if"), this will create a competition to produce better data (and to test whether the first sets of studies got their claims right).

TABLE 6. Comparison of negative citation counts for top twenty citation-receiving opinions

Panel A: Negative citation counts for the judges with the highest top twenty citation counts

Judges with Highest Top Twenty Outside-Circuit Citation Counts	(A) Outside-Circuit Citations to Top Twenty Opinions	Negative Outside Circuit Citations to Top Twenty Opinions	Negative Outside Circuit Citations as Fraction of (A)
Sandra Lynch	734	67	0.091
Frank Easterbrook	667	79	0.118
Paul Kelly	654	22	0.034
Richard Posner	570	65	0.114
Bruce Selya	516	33	0.064
Anthony Scirica	496	39	0.079
Frank Hull	460	72	0.157
Karen Williams	455	17	0.037
Edward Earl Carnes	444	38	0.086
J. Harvie Wilkinson	425	48	0.113

Panel B: Negative citation counts for eleven judges, centered on the judge with the median number of outside-circuit citations for his top twenty opinions

Median Judges	(A) Outside-Circuit Citations to Top Twenty Opinions	Negative Outside-Circuit Citations to Top Twenty Opinions	Negative Outside-Circuit Citations as Fraction of (A)
David Tatel	265	62	0.234
Edith Jones	262	20	0.076
James Loken	258	26	0.101
A. Wallace Tashima	257	25	0.097
Stephanie Seymour	256	22	0.086
James Larry Edmondson	255	42	0.165
Michael Murphy	253	24	0.095

Median Judges	(A) Outside-Circuit Citations to Top Twenty Opinions	Negative Outside-Circuit Citations to Top Twenty Opinions	Negative Outside-Circuit Citations as Fraction of (A)
Eugene Davis	250	25	0.100
Alice Batchelder	250	23	0.092
Diana Murphy	245	15	0.061
Carlos Lucero	240	13	0.054

On average, the negative citations are 0.089 of the total outside-circuit citations, (A), for the top judges. In comparison, the median judges' negative citations are 0.106 of their total outside-circuit citations. If anything, the top judges had a lower fraction of negative citations in their total citation count, inconsistent with the outrageousness hypothesis. The difference in mean fractions is not statistically significant (t=-0.833).

Another concern with the use of citation counts is that more senior judges (and the chief judge of a circuit, in particular) have the ability to assign themselves the "choice" opinions from the panels on which they serve. Thus, seniority, independent of the inherent ability of a particular judge, may drive higher citation counts. To assess the importance of seniority and chief judge status, we again compared the top ten judges based on their top twenty citation-receiving opinions against the eleven median judges. We also compared both against the bottom ten judges. (Appendix Table I reports the results.) No statistically significant difference exists in terms of seniority across the three groups of judges. Moreover, the number of chief judges in the group of top ten judges (Judge Posner and Judge Wilkinson) equals the number of chief judges in the bottom ten judges group (Judge Edwards and Judge Martin).

The tests described above do not exhaust the criticisms against the citation measures that we report.[45] Assuming, however, that these measures

45. One question that has come up at multiple workshops is whether citation counts are biased as a function of political affiliation. Perhaps Republican judges will cite other Republicans more and, if the judiciary is dominated by Republican judges, this will result in the Republicans always winning. Our response is that it would certainly be interesting to look at whether Republicans (or Democrats) are more likely to cite others from their own party than they are those from an opposing party. But, for purposes of a tournament, where presumably the President will only be picking from among those judges affiliated with his party, this in-party citation effect should not matter. After all, the President (and anyone else) is only going to look at in-party rankings. There are, however, other biases that a fuller tournament study might wish to check for, such as those on account of race, gender, and educational background.

represent a good first cut at the question, which of them should we use? We save the discussion of how to combine measures for later in the Article. For now note that across the entire tournament sample, the outside-circuit citation measure is highly correlated with the Supreme Court citations, law review and periodical citations, and top twenty opinion citation measures (see Table 7 below).

TABLE 7. Correlation matrix across the different citation-based quality measures

	Total Outside-Circuit Citations	**Total U.S. Supreme Court Citations**	**Law Review and Periodical Citations**	**Outside-Circuit Citations to Top Twenty Opinions**
Total Outside-Circuit Citations	1			
Total U.S. Supreme Court Citations	0.589	1		
Law Review and Periodical Citations	0.827	0.571	1	
Outside-Circuit Citations to Top Twenty Opinions	0.897	0.486	0.690	1

Aside from citations, another method exists of measuring quality. Standard practice is to cite an opinion without mentioning the author's name. On rare occasions, though, when the judge's name adds special significance, the judge's name will be invoked (an event we refer to as an "invocation"). So, a judge who wants to add additional authority to a point she is making might say the following: "See, Judge Fancy said much the same in her opinion." The key element is that the author is invoking the reputation of Judge Fancy to add authority to the citation. Judges with greater reputations will receive more invocations than those with lesser reputations. Invocations, therefore, provide a finer tuned measure of a

Finally, one might parse the citation data by separating the citations in diversity cases from those in cases under direct federal jurisdiction. Some judges may be better at common-law-type decisionmaking and others may be better at dealing with the more federal statutory types of questions. For those who think that it is the judges who are more skilled at the latter type of decisionmaking that are the better candidates for the Court, this parsing would provide useful information.

judge's authority and influence among her peers than do citation rates.[46] Citation rates are more about individual opinion quality, whereas invocation rates, although also attached to a citation to a specific opinion, are more of an indication of reputation.

To maintain a level contest, we focus solely on invocations to opinions from 1998 to 2000. Judge Easterbrook, for example, may well have a large number of invocations from opinions written prior to 1998. We set Judge Easterbrook's past invocation count to zero, however, and focus on how many times courts invoked Judge Easterbrook's name for opinions authored during the sample time period. Invocations are counted for each judge's majority, concurring, and dissenting opinions. Invocations are counted from all federal courts (circuit, district, and Supreme Court) as well as state courts and are measured as of May 31, 2003 using Westlaw's database. Invocations include reference to a specific judge's name (related to an opinion written during the sample time period) in both the text and a parenthetical with one exception: opinions that simply refer to a judge's dissenting or concurring opinions as "(Easterbrook, J., dissenting)" or "(Easterbrook, J., concurring)" are a function of the norm of citing the judge by name whenever a dissent or concurrence is cited. Therefore, they do not represent any display of extraordinary respect and are not counted as invocations.[47]

---

46. For other uses of invocation rates to measure judicial influence, see Mita Bhattacharya & Russell Smyth, *The Determinants of Judicial Prestige and Influence: Some Empirical Evidence from the High Court of Australia*, 30 J. LEGAL STUD. 223 (2001); David Klein & Darby Morrisroe, *The Prestige and Influcence of Individual Judges on the U.S. Courts of Appeals*, 28 J. LEGAL STUD. 371 (1999).

47. That said, it is rare for dissents and concurrences to be cited as a general matter because, by definition, they are the minority view. It could be argued, therefore, that any citation to a dissent or a concurrence signals that the cited opinion is viewed as special.

TABLE 8. Invocations to opinions published, 1998–2000 (for the ten judges with greatest total invocations)

	(A) Total Invocations	Z-Score of Normalized (A)	(B) Average Invocations per Opinion	Z-Score of Normalized (B)	(C) Percent of Invocations Attributable to Majority Opinion	(D) Circuit
Richard Posner	176	3.90**	0.65	6.68**	97.7%	7th
Frank Easterbrook	103	3.36**	0.44	4.67**	99.0%	7th
Guido Calabresi	23	1.85	0.23	2.35**	91.3%	2d
J. Harvie Wilkinson	19	1.66	0.18	1.73	73.7%	4th
Michael Boudin	13	1.30	0.10	0.70	84.6%	1st
Patrick Higginbotham	12	1.22	0.12	0.96	41.7%	5th
Diarmuid O'Scannlain	11	1.14	0.08	0.43	54.5%	9th
Edith Jones	11	1.14	0.11	0.83	81.8%	5th
Diane Wood	10	1.05	0.05	0.01	20.0%	7th
J. Michael Luttig	10	1.05	0.12	0.96	80.0%	4th

** Indicates a Z-score of 1.96 or higher (representing a two-sided probability of <5% for a normal distribution. Normalized (A) is equal to LN(1+Invocations). Normalized (B) is equal to LN(1+Average Invocations per Opinion).

Summary statistics for (A) (n=98): Mean = 6.827; Median = 3.000; Standard Deviation = 20.36; Kurtosis = 54.685; Skewness = 7.145.

Summary statistics for normalized (A) (n=98): Mean = 1.379; Median = 1.386; Standard Deviation = 0.973; Kurtosis = 2.091; Skewness = 0.750.

Summary statistics for (B) (n=98): Mean = 0.053; Median = 0.038; Standard Deviation = 0.084; Kurtosis = 31.800; Skewness = 5.085.

Summary statistics for normalized (B) (n=98): Mean = 0.049; Median = 0.037; Standard Deviation = 0.068; Kurtosis = 24.700; Skewness = 4.372.

Chi-squared test of the null hypothesis that the distribution of circuits is identical between the top judges and the bottom judges for the number of invocations (A): $\chi^2 = 13.863$ (11 d.f.) ($p \leq 0.241$). Top judges are defined as those who are in the top fifty percent of judges in the entire sample (n=98) based on the number of invocations (A). Bottom judges are defined as those who are in the bottom fifty percent of judges based on the number of invocations (A).

Table 8 (column A) reports the invocation rates for opinions written between 1998 and 2000. (Appendix Table F reports invocation rates for the entire sample of judges.)

The Posner-Easterbrook dominance on invocation numbers is dramatic. Judge Posner has at least seven times the invocations of the next highest judge below Judge Easterbrook; Judge Easterbrook has nearly five times the invocations of his closest competitor. The mean number of invocations for the sample is 6.8 with a standard deviation of 20.4. Both Judge Posner and Judge Easterbrook have at least five times the standard deviation number of invocations compared with the mean judge.

The Posner-Easterbrook dominance in invocations does not appear to be due (at least primarily) to differences in circuit norms. The chi-squared test cannot reject the null hypothesis that the distributions of circuits are identical for the top and bottom judges (divided based on the total number of invocations) ($\chi^2 = 13.863$ (11 d.f.) ($p \leq 0.241$)). Nonetheless, to correct for the effect of high opinion numbers, column (B) reports the average number of invocations per opinion written in the 1998–2000 period. The Easterbrook-Posner domination over other federal circuit judges remains.

As an aside, note the fraction of Judge Posner and Judge Easterbrook's invocations that are due to their majority opinions. For them, it is close to 100%. For almost everyone else (except Judge Guido Calabresi), a significant fraction of their invocations are to their dissents and concurrences. At the low end, for example, only twenty percent of Judge Diane Wood's invocations are to her majority opinions. Dissents (and to a lesser extent concurrences) are extraordinary events for most judges, perhaps leading the judges to devote special attention to such opinions. If one assumes such a dynamic to be at work, one might conclude that Judges Posner, Easterbrook, and Calabresi devote similar skill and attention to a large number of their majority opinions.

C. MEASURING INDEPENDENCE (AND EXTRA EFFORT)

The appointment of federal judges is a political process. Independence, and particularly independence from one's political sponsors, however, is one of the qualities that many care about in judges. It is largely for reasons of independence that we give federal judges lifetime appointments. It is also for this reason that many feel disappointment when they see Supreme Court votes fall according to the known political sympathies of the judges.[48]

48. The votes in *Bush v. Gore*, 531 U.S. 98 (2000), being a prime example of the type of voting pattern that caused great disappointment and produced heated accusations of partisanship. *See, e.g.*, Jack M. Balkin, Bush v. Gore *and the Boundary Between Law and Politics*, 110 YALE L.J. 1407 (2001); Linda Greenhouse, *Collision with Politics Risks Court's Legal Credibility*, N.Y. TIMES, Dec. 11, 2000,

A different concern about judicial independence stems from the nature of multimember courts generally. Because appointments are for life and because most courts are made up of a small number of judges, judges have to learn to get along with each other. There is the danger that they will develop the kinds of close working relationships and friendships that can deter them from openly disagreeing and antagonizing each other. In such contexts, we think, a willingness to speak independently is a trait that should be valued (regardless of the political parties of the opposing judges). More importantly, this is a trait crucial for someone on the High Court.

As measures of the willingness to disagree (or be disagreeable), we first report numbers of dissents and concurrences written (collectively referred to as "independent opinions"). The number of independent opinions captures the willingness of a judge to disagree with her colleagues on the bench. It is also an indication of the judge's willingness to displease her colleagues. We say this because we suspect that even the threat of a dissent or concurrence forces the writer of the majority opinion to exert greater effort on the opinion than what she would have exerted otherwise. The real extra effort, nonetheless, lies with the judges writing the independent opinions. These judges do not generally receive a break on the number of majority opinions they are assigned if they write more dissents and concurrences. Writing those dissents and concurrences is additional work that the judges have to do, over and above their regular load of assigned majority opinions. In sum, the number of dissents and concurrences one writes provides not only a measure of a judge's willingness to annoy their colleagues, but also a measure of their willingness to exert extra effort.

We view dissents and concurrences as a positive. We recognize, however, that some others view them as negative. For example, a high number of such independent opinions might be viewed as a sign that the writer is uncollegial or unwilling to be a team player. Alternatively, the willingness to write independently might signal that the judge has a strong ideology (one on which they are unwilling to compromise). A person with these views could simply place a negative weight on the number of independent opinions written and run the tournament that way. With transparency regarding the promotion criteria being used, people can make their own determinations about what constitutes merit.[49]

---

at A1; Mary McGrory, *Supreme Travesty of Justice*, WASH. POST, Dec. 14, 2000, at A3; Jeffrey Rosen, *The Supreme Court Commits Suicide*, NEW REPUBLIC, Dec. 25, 2000, at 18.

49. One question we considered was whether to normalize the independence data for circuit effects. Norms do appear to play a role in that independent opinions seem more common on circuits like

In our first measure of independence, we looked at numbers of dissents and concurrences regardless of who they were against. With our second measure, we attempt to factor in ideological biases.[50] Here we use data on both the dissents that judges write and the dissents that are written against them. We classify each judge's opposing opinions as either against a judge of the same party or not (and do not count dissents against unsigned opinions). Where more than one judge is on the opposing side, we focus on the opinion writer (and if multiple dissenting opinion writers exist, we treat the opinion as written by a judge of the same party if any of the opinions meet this criterion). As a proxy for the political party of each judge, we look to the party of the president who appointed the judge to the circuit court.

We then obtain the percentage of opposing opinions in which each judge opposed a judge of the same political party. Standing alone, however, such a measure is problematic. In a circuit comprised of judges of almost all the same political party, any one judge will tend to score high simply because judges of other parties are not present on the circuit. In part, this problem is countered by the presence of senior judges visiting from other circuits or district court judges sitting by designation. As a further control, for each judge we determined the political party (as proxied by the party of the appointing president) of the *other* active judges on each circuit from 1998 to 2000 (including those who eventually took senior status or retired). For judges active over the entire 1998 to 2000 period, we gave a weight of 3. For judges active for only part of the time period (for example, just for 1998) we gave a correspondingly lower weight (for example, a weight of 1).

For example, Judge Posner on the Seventh Circuit faced a mixture of judges consisting of seventy-five percent Republican-nominated judges. Even if Judge Posner (a Republican-nominated judge) simply opposed other judges regardless of political affiliation, we would expect him to oppose other Republican-nominated judges seventy-five percent of the time. Call the baseline percentage of Republican-nominated judges on the circuit the "predicted same-party opposing" rate for Republican judges.

---

the Third and Ninth. In our measures of total productivity reported in Table 3, we normalize the sum of all the opinions. In this section, however, where our focus is on the willingness to be independent, we report the raw numbers.

50. There is much literature in both political science and law that documents the effects of ideology on voting patterns in certain categories of cases. *See, e.g.*, Richard L. Revesz, *Environmental Regulation, Ideology, and the D.C. Circuit*, 83 VA. L. REV. 1717 (1997) (demonstrating in an empirical study of environmental law-related cases that judicial decisions in the D.C. Circuit are significantly correlated with the political party of the President who nominated particular judges).

To obtain our more refined measure of independence we calculate the following:

> Independence = (Actual Same Party Opposing Fraction) – (Predicted Same Party Opposing Fraction)

So, if a Republican-nominated judge on the Seventh Circuit (such as Posner) only dissented against other Republicans half of the time, she would score a -0.25 under our measure, dissenting against same-party judges twenty-five percentage points less than an independent acting judge would. The more negative the independence score, the more aligned particular judges are with their party line. On the other hand, if the Republican-nominated judge on the Seventh Circuit dissented against other Republicans ninety-five percent of the time, she would score a +0.20, being twenty percentage points more likely to dissent against a judge of the same political party. Table 9 reports the two measures of independence.[51] (Appendix Table G reports the independence measures for the entire sample of judges.)

---

51. A finer tuned measure of independence—given the focus in the current system on votes on key hot-button issues—would be to look at the proportion of votes against same-party judges on these hot-button issues. *Cf.* Cass R. Sunstein, David Schkade & Lisa Michelle Ellman, *Ideological Voting on Federal Courts of Appeals: A Preliminary Investigation* (Univ. of Chic. Law Sch. John M. Olin Law & Econ. 2d Series, Working Paper No. 198, 2003), *at* http://papers.ssm.com/abstract=442480 (reporting a higher likelihood of voting along party lines on certain key issues like abortion). We are in the process of collecting that data for a separate article.

TABLE 9. Number of opposing opinions and independence ratings, 1998–2000 (for the twenty judges with the best independence ratings)

	(A) Independence Rating	(B) Number of Dissents and Concurrences	(C) Dissents and Concurrences (Adjusted for Intercircuit Differences)	Z-Score for (C)	(D) Circuit
David Ebel	0.000	12	25	0.16	10th
R. Lanier Anderson	0.000	10	24	0.06	11th
J. Michael Luttig	-0.006	19	26	0.35	4th
Samuel Alito	-0.015	18	29	0.65	3d
Carl Stewart	-0.017	2	10	-1.49	5th
Juan Torruella	-0.018	10	27	0.44	1st
Diane Wood	-0.018	21	27	0.44	7th
Richard Posner	0.019	15	21	-0.26	7th
Edward Earl Carnes	-0.022	14	28	0.53	11th
Anthony Scirica	-0.023	2	13	-1.22	3d
Mary Schroeder	0.023	5	5	-2.12**	9th
Terrence Evans	0.024	18	24	0.09	7th
Harold DeMoss	0.024	30	38	1.78	5th
Daniel Manion	0.028	20	26	0.33	7th
Merrick Garland	-0.037	3	15	-1.01	D.C.
Frank Easterbrook	-0.042	20	26	0.33	7th
Sandra Lynch	-0.043	7	24	0.09	1st
Stephen Trott	-0.044	19	19	-0.49	9th
Joel Flaum	-0.044	10	16	-0.84	7th
Edith Jones	-0.045	21	29	0.73	5th

** Indicates a Z-score of 1.96 or higher (representing a two-sided probability of <5% for a normal distribution). The number of dissents and concurrences for each judge is adjusted so that the mean number of total dissents and concurrences for each circuit is identical and equal to 23.167 (the unadjusted mean number of total opinions for the Ninth Circuit).

Summary statistics for (A) (n=98): Mean = -0.062; Median = -0.057; Standard Deviation = 0.189; Kurtosis = 0.307; Skewness = 0.307.

Summary statistics for (B) (n=98): Mean = 14.469; Median = 13.00; Standard Deviation = 9.692; Kurtosis = 1.344; Skewness = 1.042.

Summary statistics for (C) (n=98): Mean = 23.170; Median = 22.869; Standard Deviation = 8.557; Kurtosis = 0.796; Skewness = 0.646.

Chi-squared test of the null hypothesis that the distribution of circuits is identical between the top judges and the bottom judges for the independence rating (A): $\chi^2 = 23.110$ (11 d.f.) ($p \leq 0.017$). Top judges are defined as those who are in the top fifty percent of judges in the entire sample (n=98) based on the independence rating (A). Bottom judges are defined as those who are in the bottom fifty percent of judges based on the independence rating (A).

Column A reports the judges ranked in order of absolute closeness to zero on the independence scale. A score of zero suggests no bias either against or for one's own political party of appointment. In reading the numbers in column A, however, it is important to keep in mind the total number of dissents and concurrences reported in column B. The reason we say this is that a score of zero or close to it is a lot more meaningful as a sign of independence if one is writing a lot of independent opinions than if one is not writing any.

Note from Table 9 that while Judge Posner is still in the top ten (and Judge Easterbrook in the top twenty), other judges who scored well in terms of productivity and quality are not ranked as highly in terms of independence. Judge Reinhardt, for example, wrote the second highest intercircuit adjusted number of opinions. Judge Reinhardt, however, is not among the top twenty independent judges. Also note that judges who are the most independent judges (as given in column (A)) are not the most prolific authors of dissenting and concurring opinions. The three highest dissent and concurrence opinion writers in the top twenty independent judges are in the bottom two-thirds of the list. Judge Edith Jones, for example, wrote twenty-one dissents and concurrences (tied for second highest on the list) and is the last person on the list in terms of independence ranking.[52]

To combine the opposing opinion and independence measures, we created a subsample of judges who scored between -0.100 and +0.100 on the independence measure. Such judges are not overly influenced either to

52. The chi-squared test in Table 9 rejects the null hypothesis that the distributions of circuits among the top and bottom judges (segmented based on their independence rating) are identical ($\chi^2 = 23.110$ (11 d.f.) ($p \leq 0.017$)). Because we do not control directly for intercircuit variations in independence, some of a judge's independence, therefore, may be more of a reflection of the norms within a circuit rather than the judge's own preferences. On the other hand, it is also possible that judges in a particular circuit are systematically more internally driven by ideology in their decisionmaking.

side with or against judges of the same political party. Arguably, this measure (closeness to zero on the independence scale) is a better measure of independence than a large positive score because closeness to zero suggests even-handedness—that is, it suggests neither a preference nor an aversion to side with others appointed by a president of the same party. In contrast, a high positive score, while suggesting a change in ideology, still suggests an ideologically driven position. Within this subsample, we then ranked judges based on the total number of opposing opinions authored. To control for intercircuit differences in opposing opinion authoring norms, we adjusted each judge's number of dissents and concurrences to make equal the mean number of dissents and concurrences across the circuits. Table 10 reports the results.

Note from Table 10 that Judge Harold DeMoss and Judge Rosemary Barkett score the highest in terms of number of opposing opinions. Neither of their scores, however, were statistically different from the entire sample's mean judge's score in terms of total number of dissents and concurrences.

TABLE 10. Ranking of independence among "independent judges" (-0.100 to +0.100 range of independence rating), 1998–2000

	(A) Number of Dissents and Concurrences (Adjusted for Intercircuit Differences)	**Z-Score of (A)**	**(B) Independence Rating**	**(C) Circuit**
Harold DeMoss	38	1.78	0.024	5th
Rosemary Barkett	37	1.58	0.056	11th
Morris Arnold	35	1.40	-0.060	8th
Andrew Kleinfeld	30	0.80	0.087	9th
Edith Jones	29	0.73	-0.045	5th
Samuel Alito	29	0.65	-0.015	3d
Danny Boggs	28	0.58	-0.082	6th
Edward Earl Carnes	28	0.53	-0.022	11th
Diane Wood	27	0.44	-0.018	7th
Juan Torruella	27	0.44	-0.018	1st
Marjorie Rendell	27	0.42	-0.049	3d

Summary Statistics for (A) (n=98): Mean = 23.170; Median = 22.869; Standard Deviation = 8.557; Kurtosis = 0.796; Skewness = 0.646.

IV. COMBINING THE CRITERIA

What use is a tournament if there is no winner? The answer is not obvious. To consider how to determine a winner of the tournament, we narrow the sample down to only judges less than sixty-five years old (measured as of 2003). While including all active judges in our sample allows us to rank judges based on relative performance with their peers, in practice, we suspect that the president will select a Supreme Court nominee younger than sixty-five because a younger appointee will likely have a longer tenure (and, therefore, greater influence).

A. DO WE NEED TO PICK A WINNER?

Do we need to pick a winner? One answer is that there should be multiple winners. Consider several of the basic criteria discussed above and who would win the tournament based on these criteria as reported in Table 11.

TABLE 11.

Productivity

Measure	**Top Scorer**	**Runner Up**	**3rd**	**4th**	**5th**
Number of Opinions	***Posner***	***Easterbrook***	Wood	Ripple	M. Arnold
Number of Opinions Adjusted for Circuit	***Posner***	***Easterbrook***	K. Moore	Gilman	Jerry Smith

Quality and Respect

Measure	**Top Scorer**	**Runner Up**	**3rd**	**4th**	**5th**
Total Outside Citations	***Posner***	***Easterbrook***	Lynch	Kelly	Ripple
Citations to Top Twenty Opinions	***Lynch***	***Easterbrook***	Kelly	Posner	Scirica
Total Invocations	***Posner***	***Easterbrook***	Wilkinson	Boudin	Higginbotham

Independence

Measure	Top Scorer	Runner Up	3th	4th	5th
Highest Independence Measure	***Ebel***	***Luttig***	Alito	Stewart	Wood
Most Dissents and Concurrences (for those who fall within -0.1 and +0.1 range of independence)	***Barkett***	***M. Arnold***	Kleinfeld	E. Jones	Alito

While Judge Posner and Judge Easterbrook dominate many of the measures, they are not universal in their dominance. Rather than get into the debate of which criteria to use in the tournament (or whether, indeed, other criteria exist), we could imagine a system that selects from the pool of top two finishers including Judges Posner, Easterbrook, Lynch, Ebel, Luttig, Barkett, and Arnold.

Not having one clear winner has its advantages. In the absence of any particular path to success, the possible moral hazard problems of a tournament are reduced. For example, if the tournament focused solely on opinion output, some judges might focus on publishing as many opinions as possible regardless of quality or independence. Having the president select among a pool of winners reduces the chance that judges will game the system to maximize a particular criterion.

Moreover, the benefit of a tournament may lie not so much in picking one winner as in eliminating large numbers of judges from contention for a Supreme Court position. While the president is left with some discretion, it is not limitless. A president seeking to nominate a judge who is not in the pool of winners will face pressure to justify explicitly why his candidate is the "best."[53] Such a claim is easier to make if there are no judges against whom the best claim must be measured. With a pool of tournament winners, it becomes more difficult for a president to claim that someone outside the pool is the best, thereby revealing more explicit political explanations for the selection of a particular judge.

53. In the weak form of the tournament, this pressure arises solely from public attention to the extent that the objective factors of the tournament are easily transmitted and absorbed by the public. *See supra* text accompanying notes 28–29 (discussing the weak form of the tournament).

B. COMPOSITE MEASURES

More is possible, nonetheless, than simply picking multiple winners of the tournament. One could attempt to combine the various measures of merit into a combined metric. Thus, a judge who scores best across a series of the measures would win the tournament.

As discussed above, naming several winners along multiple dimensions of merit may help reduce the benefit (and, therefore, the incentive) to focus single-mindedly on one dimension of merit. A composite measure of quality also helps reduce the possibility that judges will game the tournament in ways that detract from their performance as judges. For example, we use citation counts as a measure of quality. But citations, we know, are not perfectly correlated with quality. Judges will, therefore, have an incentive to focus on writing the type of opinions that produce more citations, and those might not be the opinions that best clarify the law and explain the resolution to the parties. Assuming that it is the longer and more complex opinions that produce more citations—a questionable assumption—there is a cost to having judges pursue citations. Such behavior uses up resources that could be used to write other opinions. To the extent that total number of opinions written is a measure of quantity of work done, and quantity is also an element of the composite measure, reducing quantity is something that the judge will be reluctant to do.[54] Hence, the cost of "bad" game-playing is unlikely to be high. That said, we concede that there might well be some game-playing of the sort described. The relevant question, however, is not whether there will be any game-playing, but whether this game-playing will be worse than that under the current system (where gaming likely already occurs along ideological lines).[55]

One could imagine a composite ranking based on the sum of the ordinal rankings of each judge across each factor of analysis (for example, productivity, quality, and independence). The simple ordinal ranking scheme, however, has at least two defects. First, it does not give any weight to the magnitude of a judge's relative performance compared to another judge. If Judge X writes ten times more opinions than the next highest judge this should get more weight than if Judge X writes only 1.1 times the

54. Interestingly, Landes, Lessig, and Solimine report in their citation study that those judges who write the most opinions also tend to score higher on average opinion quality measured through citation counts per opinion. *See* Landes et al., *supra* note 13, at 302–03.

55. *See supra* Part II (discussing the political gaming under the current ideology-driven nomination process to the Supreme Court).

number of opinions. Second, the ordinal ranking scheme gives equal weights to the productivity, quality, and independence rankings. As a matter of policy and preference, however, it is unclear why these various measures should receive equal weight.

We therefore eschew ordinal ranking schemes. Instead, we rank judges based on a composite of the productivity, quality, and independence measures. For these measures we use the following:

Productivity: Number of opinions (majority, dissenting, and concurring) from 1998 to 2000 adjusted for intercircuit differences.

Quality: Equal weight linear combination of

(a) the natural log of the outside citations for the top twenty opinions

and

(b) the natural log of (1 + the number of invocations) (multiplied by a constant to equalize the standard deviation with (a)).

While other weightings are possible, the high correlation between terms (a) and (b) makes it unlikely that this will make a significant difference.

Independence: The absolute value of the independence measure (where the farther a judge is away from zero, the more the judge is affected by the political party of the opposing judge).

To construct the composite, we multiply each measure by a constant to equalize the standard deviations across the measures. Equalizing the standard deviations gives roughly the same weight to variations in each variable in generating a composite measure. The means for the productivity and quality measures are then normalized to zero.

We then provide a series of weighting for each measure in a linear combination as follows:

$$\text{Score} = x\text{Productivity} + y\text{Quality} - z\text{Independence}$$

Such that

$$x + y + z = 1$$

Table 12 reports the results of select weightings for all judges in the sample excluding those sixty-five and over. Before we proceed, we have a few caveats.

1. The Superstar Effect

We treat judicial opinions as products and then use citation and invocation rates as measures of customer use of the products. Judicial opinions, however, are unusual products in that they all have the same cost. Indeed, for judges and academics, because of either government-paid or free access to the Westlaw and Lexis databases, the cost is zero. This means that if one opinion is even slightly better than another one, everyone will use the better one.[56] And if there are network effects—where, for example, the authority of an opinion increases with the number of citations it receives—the incentive to pick the slightly better opinion increases further. The result, then, is a situation where all the citations go to the better opinion even if it is only slightly better than the next best one. So, if opinion A receives 100 times the citations that opinion B (on the same topic) does, that does not mean opinion A is 100 times better than opinion B. In tournament terms, this means that one might wish to discount a high relative score on the measures, such as citations and invocations, that make up our quality measure and are subject to superstar effects. In part, our use of log transformations of the number of citations to the top twenty opinions and of the number of invocations helps reduce the skewness in the distribution due to the superstar effect (resulting in a more normal distribution of judges).

2. Small Numbers of Dissents

The second of our independence measures looks at numbers of disagreements with judges from the same political party. On some circuits, however, judges dissent very little and this can create a problem. So, for example, Judge Carl Stewart receives a high independence score. He, however, wrote only two dissenting and concurring opinions in the three-year period. Judge Stewart's high independence score could therefore simply be an artifact of his low number of dissents. For ease of computation, nonetheless, we use only the simple independence ranking without regard to a judge's output of dissenting and concurring opinions.

56. The classic papers on superstar theory include Moshe Adler, *Stardom and Talent*, 75 AM. ECON. REV. 208 (1985); Glenn M. MacDonald, *The Economics of Rising Stars*, 78 AM. ECON. REV. 155 (1988); Sherwin Rosen, *The Economics of Superstars*, 71 AM. ECON. REV. 845 (1981).

More refined versions of the tournament may wish to take into account opposing opinion productivity in addition to political independence.

TABLE 12. Linear combinations of quality, productivity, and independence

Measures for judges under sixty-five with various weightings

(quality, productivity, independence)

[Quality]

Posner
Easterbrook
(1, 0, 0)

Posner Easterbrook (0.75, 0.25, 0) | **Posner** Easterbrook (0.75, 0, 0.25)

Posner
Easterbrook
(0.67, 0.16, 0.16)

Posner Easterbrook (0.5, 0.5, 0) | **Posner** Easterbrook (0.5, 0, 0.5)

Posner Easterbrook (0.42, 0.42, 0.16) | **Posner** Easterbrook (0.42, 0.16, 0.42)

Posner
Easterbrook
(0.33, 0.33, 0.33)

Posner Easterbrook (0.25, 0.75, 0) | **Posner** Easterbrook (0.16, 0.67, 0.16) | **Posner** Easterbrook (0.16, 0.16, 0.67) | **Posner** Easterbrook (0.25, 0, 0.75)

Posner
Easterbrook
(0.16, 0.42, 0.42)

Posner Easterbrook (0, 1, 0)	**Posner** Easterbrook (0, 0.75, 0.25)	**Posner** Easterbrook (0, 0.5, 0.5)	**Posner** Ebel (0, 0.25, 0.75)	**Ebel** Luttig (0, 0, 1)

[Productivity] **[Independence]**

Critics have questioned our measures on the grounds that we have no evidence they predict high-quality Supreme Court performance. The implicit suggestion is that we must first do an analysis of relative performance for current (and past) Justices on the Supreme Court and then work backward to see which characteristics of the circuit court judges (or state court judges or others) best predict Supreme Court performance. The idea is an attractive one.[57] Our initial thinking about the suggestion is that such an analysis would be difficult, although probably not impossible. It is

57. The one study that we are familiar with that does use objective data to make relative performance comparisons among the Justices is Lee Epstein et al., *supra* note 10. This study does not, however, use the rankings to come up with a set of criteria that could help predict performance.

difficult because, among other things, there are only nine Justices at any one time, creating a problem of adequate numbers. We can imagine a study, however, that adjusts for differences in caseloads and such and compares Justices across the decades.[58]

There are dozens of other caveats we could add. These measures are imperfect proxies by nature. But even these imperfect proxies, especially if one looks at enough of them, contain information.

The striking result from the multiple weightings of our composite measure of quality reported is the dominance of Judge Posner across a wide range of weights reported in Table 12. In the rankings, Judge Posner wins in every possible ranking except the one in which independence is given 100% of the weight. There, Judge Ebel wins instead. Judge Easterbrook likewise wins nearly as many second-place finishes. People may care about a number of different criteria, but the conflict among criteria may be more illusion than reality. At least in the case of Judge Posner, one judge does seem to dominate across a range of different measures, regardless of the weighting.

Of course, in conducting a tournament of judges over time, there may not be a judge like Judge Posner. The almost near domination of Judge Easterbrook as the second place judge, however, stands as a counterexample to this possibility. Nonetheless, to examine how the tournament would play out without Judge Posner and Judge Easterbrook, we took them out and ranked judges according to the same range of weightings used for Table 12 on the quality, productivity, and independence measures. The results are reported in Table 13.

---

58. Steve Goldberg suggested, interestingly, that our data could be used to work backward to see which background characteristics correlated with high performance on the circuit courts. That information, he suggested, could then be used to construct a tournament for positions on the circuit courts.

TABLE 13. Linear combinations of quality, productivity, and independence

Measures for judges under sixty-five with various weightings, excluding Judges Posner and Easterbrook

(quality, productivity, independence)

[Quality]

S. Lynch
Wilkinson
(1, 0, 0)

Wilkinson
S. Lynch
(0.75, 0.25, 0)

S. Lynch
Wilkinson
(0.75, 0, 0.25)

Wilkinson
S. Lynch
(0.67, 0.16, 0.16)

Wilkinson
K. Moore
(0.5, 0.5, 0)

S. Lynch
Luttig
(0.5, 0, 0.5)

Wilkinson
Niemeyer
(0.42, 0.42, 0.16)

S. Lynch
Wilkinson
(0.42, 0.16, 0.42)

Wilkinson
Niemeyer
(0.33, 0.33, 0.33)

K. Moore
Gilman
(0.25, 0.75, 0)

K. Moore
J. Smith
(0.16, 0.67, 0.16)

Ebel
Luttig
(0.16, 0.16, 0.67)

S. Lynch
Luttig
(0.25, 0, 0.75)

Ebel
J. Smith
(0.16, 0.42, 0.42)

K. Moore	**J. Smith**	**Ebel**	**Ebel**	**Ebel**
Gilman	K. Moore	J. Smith	Carnes	Luttig
(0, 1, 0)	(0, 0.75, 0.25)	(0, 0.5, 0.5)	(0, 0.25, 0.75)	(0, 0, 1)

[Productivity] **[Independence]**

Without Judge Posner or Judge Easterbrook, no one winner emerges. Several judges, however, do come to the forefront, including most notably Judge Wilkinson and Judge Lynch (with the two highest numbers of first and second place finishes). As with conducting multiple tournaments, the combination of the various criteria could be used to select multiple winners. Moreover, Table 13 allows observers to see the trade-offs among the criteria.

C. AND WHAT ABOUT THE LOW SCORERS?

Pointing out the low scorers may give circuit court judges an incentive to produce more and higher quality opinions (while avoiding simply toeing the party line). Additionally, identifying low scorers may help eliminate

nominees that the president may otherwise be tempted to put forth to the Senate. Or, alternatively, it may give the Senate more fodder with which to critique a nominee.

We are not able to say much about the low scorers, however. Unlike with the high scorers' results, in which a few judges dominated, the scores in the bottom half are tightly clustered. The tight clustering means that the names on the list of those at the bottom will change significantly if we alter the weights on the different criteria even slightly. This is in contrast to Table 12, in which Judge Posner and Judge Easterbrook dominate across the different combinations of weights. Table 14 reports the twenty lowest scoring judges using an equal weighting of the quality, productivity, and independence measures. Column B reports the worst ordinal ranking each judge received across the three measures separately. (Appendix Table H ranks all the judges based on the equal weighted composite measure.)

TABLE 14. Twenty lowest scoring judges (under sixty-five) using equal weighting for quality, productivity, and independence

Twenty Lowest Scoring Judges	**(A) Rank Based on Equal Weight Composite Measure (Lowest = 1)**	**(B) Worst Ordinal Ranking (Among the Three Measures) (Lowest = 98)**	**(C) Years as a Circuit Court Judge (in 1998)**	**(D) Circuit**
Stanley Marcus	1	98[i]	1	11th
Pamela Ann Rymer	2	96[q]	9	9th
Martha Daughtrey	3	95[p]	5	6th
Michael Daly Hawkins	4	93[q]	4	9th
Alice Batchelder	5	92[i]	7	6th
Karen Henderson	6	98[q]	8	D.C.
Robert Henry	7	91[p]	4	10th
Rhesa Barksdale	8	85[p]	8	5th
M. Blane Michael	9	90[i]	5	4th
Carlos Lucero	10	94[i]	3	10th
Deanell Reece Tacha	11	93[i]	13	10th
Joel Fredrick Dubina	12	91[q]	8	11th
Susan Harrell Black	13	78[q]	6	11th
James Loken	14	97[i]	8	8th

Twenty Lowest Scoring Judges	(A) Rank Based on Equal Weight Composite Measure (Lowest = 1)	(B) Worst Ordinal Ranking (Among the Three Measures) (Lowest = 98)	(C) Years as a Circuit Court Judge (in 1998)	(D) Circuit
R. Guy Cole	15	83[i]	3	6th
David Bryan Sentelle	16	94[q]	11	D.C.
Merrick Garland	17	95[q]	1	D.C.
Alex Kozinski	18	86[p]	13	9th
Jose Cabranes	19	89[i]	4	2d
Daniel Manion	20	98[p]	12	7th

[q] Indicates that quality was the judge's worst ranking criterion; [p] indicates that productivity was the judge's worst ranking criterion; [i] indicates that independence was the judge's worst ranking criterion.

Equal Weighted Composite Measure = 0.333Quality + 0.333Productivity – 0.333Independence.

D. THE EFFECT OF EXPERIENCE

One criticism of our methodology may be that it gives undue weight to those judges with more experience on the bench. In comparing the twenty lowest scoring judges with the judges who performed the best in the tournament, however, it is unclear whether there is, in fact, an experience gap. Table 15 reports the years of experience for those judges who were one of the top two winners of each individual measure of merit, as well as those receiving either a first or second finish in the composite measures (excluding Judge Easterbrook or Judge Posner) as reported in Table 13:

TABLE 15. Years of experience for the top judges in the tournament

Top Judges in the Tournament	Years as a Circuit Court Judge (at the start of 1998)
Barkett	4
Carnes	6
Ebel	11
Gilman	1
Smith	7
K. Moore	3
Luttig	7

Top Judges in the Tournament	Years as a Circuit Court Judge (at the start of 1998)
M. Arnold	6
Niemeyer	8
Lynch	3
Wilkinson	14

The mean experience for the bottom twenty judges is equal to 6.7 years at the start of 1998 (the beginning of the tournament). The mean experience for the top judges in the tournament is equal to 6.4. The unpaired t-statistic for the difference in the mean for the top judges compared with the nontop judges is not statistically significant from zero (t = -0.205). Even when Judge Posner and Judge Easterbrook are added to the top judges, the mean experience for the top judges at the start of the tournament is only 7.7 years. The unpaired t-statistic for the difference in means is again insignificant (t = 0.667).

We suspect that a fuller study might reveal there is something of an age-experience life cycle profile, in which performance is initially low and then increases with experience and age—but then there comes a point where it begins to decrease.[59] The presence of a number of relatively junior judges among the top performers suggests that even if initial performance is low, the learning curve is steep.

E. MORE ON INTERCIRCUIT DIFFERENCES

Two of the criteria used to construct the composite measure, quality and productivity, are chosen based on measures without significant intercircuit variation. Quality is based on the outside-circuit citations to the judges' top twenty citation-receiving decisions. As discussed above, this measure controls for differences in publication norms across circuits (by looking at the same number of opinions, twenty, for each judge).[60] Similarly, productivity is based on the total number of published opinions

59. For a study along these lines (that uses citations as a measure of the quality), see Russell Smyth & Mita Bhattacharya, *How Fast do Old Judges Slow Down? A Life Cycle Study of Aging and Productivity in the Federal Court of Australia*, 23 INT'L REV. L. & ECON. 141 (2003).

60. *See supra* Part III.B.

adjusted to eliminate intercircuit variations.[61] Some intercircuit variation, nonetheless, exists in the third criterion, independence.[62]

Table 16 reports the circuit distribution of top and bottom judges, where the judges are divided based on whether they are in the top or bottom half of the rankings determined through the equal weighted composite measure.

TABLE 16. Circuit breakdown of highest and lowest scoring judges (based on the equal weighted composite measure of merit)

Circuit	Number of Top Judges	Percent of Circuit Participating in Tournament	Number of Bottom Judges	Percent of Circuit Participating in Tournament
1st	2	100.0	0	0.0
2d	2	50.0	2	50.0
3d	5	100.0	0	0.0
4th	5	71.4	2	28.6
5th	6	75.0	2	25.0
6th	4	57.1	3	42.9
7th	4	50.0	4	50.0
8th	1	50.0	1	50.0
9th	2	28.6	5	71.4
10th	3	37.5	5	62.5
11th	3	37.5	5	62.5
D.C.	0	0.0	8	100.0
Total	37		37	

Equal weighted composite measure = 0.333Quality + 0.333Productivity + 0.333Independence. Top judges are defined as those who score in the top fifty percent of judges based on the equal weighted composite measure of merit. Bottom judges are defined as those who score in the bottom fifty percent of judges based on the equal weighted composite measure of merit.

Chi-squared test of the null hypothesis that the distribution of circuits is identical between the top judges and the bottom judges (the bottom 50%): $\chi^2 = 20.714$ (11 d.f.) ($p \leq 0.037$).

Note from Table 16 that some circuits do better than others. The First and Third Circuits both have all their judges participating in the tournament in the top fifty percent of judges. Interestingly, the Seventh Circuit (home to Judge Posner and Judge Easterbrook) only has half of its judges in the

61. *See supra* text accompanying notes 36–37.

62. *See supra* Part III.C.

top half of judges. The D.C. Circuit, in contrast, performs surprisingly poorly, with all of its judges in the bottom half of judges.

On the one hand, the poor showing of certain circuits (including the D.C. Circuit) may be due to circuit court norms or institutional characteristics not inherent to a specific judge that we have failed to capture. The D.C. Circuit, for example, tends primarily to hear specialized cases (dealing with administrative law issues for example) that may both require greater effort (leading to a lower productivity) and lead to a lower outside-circuit citation count (due to the specialized circuit-specific nature of the cases).[63] On the other hand, our measures control for intercircuit differences in overall productivity. The relatively low composite scores of the D.C. Circuit judges, therefore, exist at least partly because no single judge stands out in terms of productivity relative to her peers on the circuit. While being average is not necessarily a bad thing, it is perhaps not the quality we seek in a Supreme Court Justice. In addition, the specialized nature of cases in a particular circuit should not necessarily affect how judges fare on our independence measure. The low independence scores among D.C. Circuit judges suggest a tendency to make decisions in a more ideological manner than in other circuits (thus generating a lower independence rating for the judges). Such judges may bring with them this heightened tendency if put on the Supreme Court.

While judges from certain circuits do poorly in our tournament, we are hesitant to introduce intercircuit controls beyond those we have already employed. Further, although this question requires greater attention, our results may call into question why the D.C. Circuit is often viewed as a fertile ground for future Supreme Court Justices.[64]

V. COMPARISON WITH THE BUSH "FIVE"

The rumor mill has five current federal circuit judges on President Bush's short list: Judges Samuel Alito, Emilio Garza, Edith Jones, J. Michael Luttig, and J. Harvie Wilkinson (the "Bush Five").[65] The question

63. *See* Landes et al., *supra* note 13, at 303.

64. We are grateful to Michael Solimine for flagging this issue for us.

65. *See, e.g.*, John A. MacDonald, *Highest Court, Higher Stakes Prospect of Vacancy Fuels Speculation About Supreme Court's Direction*, THE HARTFORD COURANT, May 31, 2003, at A1, *available at* 2003 WL 55332005; Stuart Taylor Jr., *Bush and the Supreme Court: Place Your Bets*, NAT'L L.J., Nov. 16, 2002, *available at* 2002 WL 26794497.

In addition to the five names mentioned in the text, two other names that are often mentioned are Alberto Gonzales, White House Counsel and former Texas Supreme Court Justice, and Justice Janice Rogers Brown of the California Supreme Court. A fuller study could compare the relative performances of the state court justices to see how Gonzalez and Brown stand against their peers as well as federal

one might have—given the opaque nature of the current promotion system—is whether the President and his advisers use anything akin to the system we propose in making their short list. Alternatively, one might ask whether the numbers provide any insight into the characteristics of those judges on the short list.

Our expectation was that the Bush Five would fare abysmally in the tournament (given our perception about the ideological focus of the current administration). Surprisingly, those in the Bush Five did well. Placing Judges Posner and Easterbrook to the side, Judges Wilkinson and Luttig emerge as among the top performers (along with Judges Lynch, Karen Nelson Moore, Ebel, and Jerry Smith). More surprising was to see that three of the Bush Five (Judges Luttig, Alito, and Jones) have among the highest scores on the independence measure.

VI. CONCLUSION

What do the numbers from the tournament tell us? Should Judge Posner and Judge Easterbrook be the president's nominees? Not necessarily. There may be finer tuned measures that do not end with them as tournament winners. More importantly, an analysis of the substance of their writings and decisions might reveal them to be either too ideological or plain crazy. Our primary goal is not to produce winners, but to enable transparency in the nomination process. Judge Posner and Judge Easterbrook are useful in this analysis in that their numbers give us a basis to challenge the president if he does not choose them. Ideally, the president and the Senate would be able to answer with a critique of our measures and an explanation of why, under a more detailed examination of the underlying data, someone else should emerge as the leader.

The tournament adds value in creating a de facto presumption—one that the president has to rebut (or else face public pressure to the extent that the tournament's objective winners are easy to observe). The choice of the nominee is that of the president with the advice and consent of the Senate.

---

circuit court judges. If they do not compare favorably, that should be reason to question claims of merit that are made regarding them. As an aside, it is noteworthy that Justice Brown has been nominated to the D.C. Circuit, supposedly, in the minds of many, as a first step toward moving her to the U.S. Supreme Court. *See* Carolyn Lochhead, *Democrats Intend to Block Vote on Bush Pick; Feinstein Opposes Nomination of State Justice to Appeals Court*, S.F. CHRON., Nov. 7, 2003, at A6; Jim Puzzanghera, *U.S. Senate Showdown Set on Stalled Judicial Nominees*, SAN JOSE MERCURY NEWS, Nov. 7, 2003, at News 1. While this Article was in the editing process, President Bush nominated Alberto Gonzales to be the U.S. Attorney General. *See* CNN.com, *Bush Attorney General Pick is Alberto Gonzales*, *at* http://www.cnn.com/2004/ALLPOLITICS/11/10/bush.cabinet/ (Nov. 11, 2004).

But the president has to justify his choice to the public. If he introduces someone other than Judge Posner or Judge Easterbrook as the most meritorious, the media needs to ask why. If asked, the president will have to explain the metric of merit he is using. And if he has no merit-based justification, the results from running a tournament of judges will force him to explain what other criteria he used. If the president's reason for picking a candidate is that she will vote to overturn *Roe v. Wade*, or that the candidate was a good campaign contributor or an old family friend, the objective should be to uncover such motives. On the flip side, the president will also have to explain why, if he claims merit matters, Judge Posner or Judge Easterbrook was not his candidate. Maybe Judge Posner is too unpredictable and will not toe the Republican party line. But independent thinking is supposed to be a qualification, not a disqualification. Or maybe the president's reason for not nominating Judge Posner is a fear that the Democrats will block the nomination on ideological grounds? If so, maybe the tournament results will induce him to nominate Judge Posner and the burden of justification will move to the Democratic senators to show why Judge Posner's ideology is problematic enough to justify blocking him.

This Article began out of our frustration with the current state of the judicial appointments process. As a result, the goal for our tournament was to provide an improvement over the current system. The harder question is whether our measures could play a role in a more bipartisan, but subjective, selection process of the type that some would claim we have had in the past. Our view is that they should; objective measures will serve as a check on the inevitable biases that any system of subjective analysis will possess (and vice versa). But that discussion is for another day and another paper.

VII. APPENDIX

TABLE A. Federal circuit court judges in the tournament

Judge	(A) Circuit	(B) Year Appointed	(C) Age in 2003*
Juan Torruella	1st	1984	70
Bruce Selya	1st	1986	69
Michael Boudin	1st	1992	64
Sandra Lynch	1st	1995	57
John Walker	2d	1989	63
Dennis Jacobs	2d	1992	65
Guido Calabresi	2d	1994	71
Jose Cabranes	2d	1994	63
Fred Parker	2d	1994	65
Dolores Sloviter	3d	1979	71
Anthony Scirica	3d	1987	63
Richard Nygaard	3d	1988	63
Samuel Alito	3d	1990	53
Jane Roth	3d	1991	68
Theodore McKee	3d	1994	56
Marjorie Rendell	3d	1997	56
H. Emory Widener	4th	1972	80
J. Harvie Wilkinson	4th	1984	59
William Wilkins	4th	1986	61
Paul Niemeyer	4th	1990	62
J. Michael Luttig	4th	1991	49
Karen Williams	4th	1992	52
M. Blane Michael	4th	1993	60
Diana Gribbon Motz	4th	1994	60
Carolyn Dineen King	5th	1979	65
Patrick Higginbotham	5th	1982	65
E. Grady Jolly	5th	1982	66
W. Eugene Davis	5th	1983	67
Edith Jones	5th	1985	54
Jerry Smith	5th	1987	57
Jacques Wiener	5th	1990	69

Judge	(A) Circuit	(B) Year Appointed	(C) Age in 2003*
Rhesa Barksdale	5th	1990	59
Harold DeMoss	5th	1991	73
Emilio Garza	5th	1991	56
Carl Stewart	5th	1994	53
Fortunato Benavides	5th	1994	56
James Dennis	5th	1995	67
Boyce Martin	6th	1979	68
Danny Boggs	6th	1986	59
Alice Batchelder	6th	1991	59
Martha Craig Daughtrey	6th	1993	61
Karen Nelson Moore	6th	1995	55
R. Guy Cole	6th	1995	52
Ronald Lee Gilman	6th	1997	61
Eric Clay	6th	1997	55
Richard Posner	7th	1981	64
John Coffey	7th	1982	81
Joel Flaum	7th	1983	67
Frank Easterbrook	7th	1985	55
Kenneth Ripple	7th	1985	60
Daniel Manion	7th	1986	61
Michael Kanne	7th	1987	65
Ilana Rovner	7th	1992	65
Diane Wood	7th	1995	53
Terrence Evans	7th	1995	63
Pasco Bowman	8th	1983	70
Roger Wollman	8th	1985	69
James Loken	8th	1990	63
Morris Arnold	8th	1992	62
Diana Murphy	8th	1994	69
Harry Pregerson	9th	1979	80
Mary Schroeder	9th	1979	63
Stephen Reinhardt	9th	1980	72
Alex Kozinski	9th	1985	53
Diarmuid O'Scannlain	9th	1986	66
Stephen Trott	9th	1988	64
Pamela Ann Rymer	9th	1989	62

Judge	(A) Circuit	(B) Year Appointed	(C) Age in 2003*
Thomas Nelson	9th	1990	67
Andrew Kleinfeld	9th	1991	58
Michael Daly Hawkins	9th	1994	58
A. Wallace Tashima	9th	1996	69
Sidney Thomas	9th	1996	50
Stephanie Seymour	10th	1979	63
Deanell Reece Tacha	10th	1985	57
David Ebel	10th	1987	63
Paul Kelly	10th	1992	63
Robert Henry	10th	1994	50
Michael Murphy	10th	1995	56
Mary Beck Briscoe	10th	1995	56
Carlos Lucero	10th	1995	63
Gerald Bard Tjoflat	11th	1975	74
Robert Anderson	11th	1979	67
James Edmondson	11th	1986	56
Stanley Birch	11th	1990	58
Joel Fredrick Dubina	11th	1990	56
Edward Earl Carnes	11th	1992	53
Susan Harrell Black	11th	1992	60
Rosemary Barkett	11th	1994	64
Frank Hull	11th	1997	55
Stanley Marcus	11th	1997	57
Harry Edwards	D.C.	1980	63
Douglas Ginsburg	D.C.	1986	57
David Bryan Sentelle	D.C.	1987	60
Arthur Randolph	D.C.	1990	60
Karen Henderson	D.C.	1990	59
David Tatel	D.C.	1994	61
Judith Wilson Rogers	D.C.	1994	64
Merrick Garland	D.C.	1997	51

* Age in 2003 is calculated by subtracting the year of the judge's birth from 2003.

TABLE B. Published opinions, 1998–2000

Judge	(A) Total Number of Published Opinions	(B) Total Number of Published Majority Opinions	(C) Circuit
Richard Posner	269	254	7th
Frank Easterbrook	233	213	7th
Joel Flaum	202	192	7th
Diane Wood	194	173	7th
Kenneth Ripple	182	151	7th
Michael Kanne	177	176	7th
Morris Arnold	175	152	8th
John Coffey	168	162	7th
James Loken	167	147	8th
Roger Wollman	158	154	8th
Terrence Evans	153	135	7th
Ilana Rovner	152	122	7th
Bruce Selya	149	145	1st
Stephen Reinhardt	142	94	9th
Pasco Bowman	140	137	8th
Diarmuid O'Scannlain	139	94	9th
Juan Torruella	138	128	1st
Michael Boudin	135	131	1st
Jerry Smith	132	118	5th
Karen Nelson Moore	130	94	6th
Ronald Lee Gilman	124	99	6th
Daniel Manion	122	102	7th
Sandra Lynch	120	113	1st
David Ebel	114	102	10th
Paul Niemeyer	113	92	4th
Danny Boggs	113	89	6th
Diana Murphy	111	106	8th
Dennis Jacobs	109	92	2d
Gerald Bard Tjoflat	108	98	11th
Harold DeMoss	108	78	5th
Emilio Garza	106	75	5th
Michael Murphy	106	102	10th
Stephen Trott	105	86	9th

Judge	(A) Total Number of Published Opinions	(B) Total Number of Published Majority Opinions	(C) Circuit
Carl Stewart	104	102	5th
J. Harvie Wilkinson	103	86	4th
Guido Calabresi	101	84	2d
Edith Jones	101	80	5th
Harry Pregerson	99	82	9th
Wallace Tashima	99	76	9th
Deanell Reece Tacha	99	92	10th
Patrick Higginbotham	97	94	5th
John Walker	95	91	2d
Fortunato Benavides	94	83	5th
Stanley Birch	92	82	11th
Dolores Sloviter	92	80	3d
Paul Kelly	92	81	10th
James Dennis	89	49	5th
Rosemary Barkett	88	65	11th
Edward Earl Carnes	86	72	11th
Carolyn Dineen King	85	75	5th
Sidney Thomas	85	67	9th
Mary Beck Briscoe	85	59	10th
Jacques Wiener	83	75	5th
Carlos Lucero	83	70	10th
Andrew Kleinfeld	82	52	9th
David Tatel	82	65	D.C.
Grady Jolly	82	72	5th
J. Michael Luttig	81	62	4th
Eric Clay	81	62	6th
Jose Cabranes	81	71	2d
Judith Rogers	79	73	D.C.
Douglas Ginsburg	79	77	D.C.
Michael Daly Hawkins	78	60	9th
Arthur Randolph	77	61	D.C.
David Bryan Sentelle	77	60	D.C.
Diana Gribbon Motz	76	62	4th
Karen Henderson	76	51	D.C.
Mary Schroeder	73	68	9th

Judge	(A) Total Number of Published Opinions	(B) Total Number of Published Majority Opinions	(C) Circuit
Harry Edwards	71	64	D.C.
H. Emory Widener	71	51	4th
Rhesa Barksdale	71	60	5th
Samuel Alito	70	52	3d
W. Eugene Davis	70	69	5th
R. Guy Cole	69	55	6th
Stephanie Seymour	68	63	10th
Alex Kozinski	67	40	9th
Pamela Ann Rymer	67	45	9th
Alice Batchelder	67	48	6th
Fred Parker	66	58	2d
Robert Henry	63	56	10th
Susan Harrell Black	59	49	11th
Karen Williams	59	50	4th
James Edmondson	58	51	11th
William Wilkins	58	47	4th
Richard Nygaard	57	45	3d
Marjorie Rendell	55	39	3d
M. Blane Michael	55	38	4th
Theodore McKee	55	41	3d
Merrick Garland	55	52	D.C.
Jane Roth	54	41	3d
Frank Hull	54	48	11th
Robert Anderson	54	44	11th
Stanley Marcus	50	49	11th
Boyce Martin	48	41	6th
Martha Daughtrey	46	38	6th
Thomas Nelson	45	39	9th
Joel Fredrick Dubina	44	40	11th
Anthony Scirica	38	36	3d

Summary Statistics for (A) (n=98): Mean = 98.1; Median = 85.5; Standard Deviation = 42.8;Kurtosis = 2.501; Skewness = 1.418.

Summary Statistics for (B) (n=98): Mean = 83.6; Median = 74.0; Standard Deviation = 41.5; Kurtosis = 2.918; Skewness = 1.576.

Chi-squared test of the null hypothesis that the distribution of circuits is identical between the top judges and the bottom judges for the number of published opinions (A): $\chi^2 = 34.697$ (11 d.f.) ($p \leq 0.0003$). Top judges are defined as those who are in the top fifty percent of judges based on the number of published opinions (majority, concurrences, and dissents) (A). Bottom judges are defined as those who are in the bottom fifty percent of judges based on the number of published opinions (majority, concurrences, and dissents) (A).

Table C. Published opinions, 1998–2000 (adjusted for intercircuit differences)

Judge	(A) Total Number of Published Opinions	Z-Score of (A)	(B) Circuit
Richard Posner	269.0	3.60**	7th
Stephen Reinhardt	237.1	2.23**	9th
Diarmuid O'Scannlain	234.1	2.10**	9th
Frank Easterbrook	233.0	2.05**	7th
Karen Nelson Moore	230.5	1.94*	6th
Ronald Lee Gilman	224.5	1.69*	6th
Gerald Bard Tjoflat	223.9	1.66*	11th
Jerry Smith	223.2	1.63	5th
Paul Niemeyer	221.2	1.54	4th
Dolores Sloviter	217.1	1.37	3d
Danny Boggs	213.5	1.21	6th
J. Harvie Wilkinson	211.2	1.12	4th
David Ebel	210.5	1.09	10th
Morris Arnold	210.0	1.06	8th
Stanley Birch	207.9	0.97	11th
Rosemary Barkett	203.9	0.80	11th
Dennis Jacobs	203.8	0.80	2d
Michael Murphy	202.5	0.74	10th
Joel Flaum	202.0	0.72	7th
James Loken	202.0	0.72	8th
Edward Earl Carnes	201.9	0.72	11th
Stephen Trott	200.1	0.64	9th
Harold DeMoss	199.2	0.60	5th
Bruce Selya	198.7	0.58	1st
Emilio Garza	197.2	0.51	5th
Guido Calabresi	195.8	0.45	2d
Deanell Reece Tacha	195.5	0.44	10th
Carl Stewart	195.2	0.43	5th
Samuel Alito	195.1	0.42	3d
Harry Pregerson	194.1	0.38	9th
A. Wallace Tashima	194.1	0.38	9th

Judge	(A) Total Number of Published Opinions	Z-Score of (A)	(B) Circuit
Diane Wood	194.0	0.38	7th
Roger Wollman	193.0	0.33	8th
David Tatel	192.7	0.32	D.C.
Edith Jones	192.2	0.30	5th
John Walker	189.8	0.20	2d
Douglas Ginsburg	189.7	0.19	D.C.
Judith Rogers	189.7	0.19	D.C.
J. Michael Luttig	189.2	0.17	4th
Paul Kelly	188.5	0.14	10th
Patrick Higginbotham	188.2	0.13	5th
Juan Torruella	187.7	0.11	1st
Arthur Randolph	187.7	0.11	D.C.
David Bryan Sentelle	187.7	0.11	D.C.
Karen Henderson	186.7	0.06	D.C.
Fortunato Benavides	185.2	0.00	5th
Michael Boudin	184.7	-0.02	1st
Diana Gribbon Motz	184.2	-0.04	4th
Richard Nygaard	182.1	-0.13	3d
Kenneth Ripple	182.0	-0.14	7th
Harry Edwards	181.7	-0.15	D.C.
Eric Clay	181.5	-0.16	6th
Mary Beck Briscoe	181.5	-0.16	10th
James Dennis	180.2	-0.22	5th
Sidney Thomas	180.1	-0.22	9th
Theodore McKee	180.1	-0.22	3d
Marjorie Rendell	180.1	-0.22	3d
Carlos Lucero	179.5	-0.25	10th
H. Emory Widener	179.2	-0.26	4th
Jane Roth	179.1	-0.26	3d
Andrew Kleinfeld	177.1	-0.35	9th
Michael Kanne	177.0	-0.35	7th
Carolyn Dineen King	176.2	-0.39	5th
Jose Cabranes	175.8	-0.40	2d
Pasco Bowman	175.0	-0.44	8th

Judge	(A) Total Number of Published Opinions	Z-Score of (A)	(B) Circuit
Susan Harrell Black	174.9	-0.44	11th
Jacques Wiener	174.2	-0.47	5th
James Edmondson	173.9	-0.49	11th
E. Grady Jolly	173.2	-0.52	5th
Michael Daly Hawkins	173.1	-0.52	9th
Frank Hull	169.9	-0.66	11th
Robert Anderson	169.9	-0.66	11th
Sandra Lynch	169.7	-0.67	1st
R. Guy Cole	169.5	-0.67	6th
Mary Schroeder	168.1	-0.73	9th
John Coffey	168.0	-0.74	7th
Alice Batchelder	167.5	-0.76	6th
Karen Williams	167.2	-0.77	4th
William Wilkins	166.2	-0.82	4th
Stanley Marcus	165.9	-0.83	11th
Merrick Garland	165.7	-0.84	D.C.
Stephanie Seymour	164.5	-0.89	10th
M. Blane Michael	163.2	-0.94	4th
Anthony Scirica	163.1	-0.95	3d
Rhesa Barksdale	162.2	-0.99	5th
Alex Kozinski	162.1	-0.99	9th
Pamela Ann Rymer	162.1	-0.99	9th
W. Eugene Davis	161.2	-1.03	5th
Fred Parker	160.8	-1.05	2d
Joel Fredrick Dubina	159.9	-1.09	11th
Robert Henry	159.5	-1.10	10th
Terrence Evans	153.0	-1.38	7th
Ilana Rovner	152.0	-1.43	7th
Boyce Martin	148.5	-1.58	6th
Martha Craig Daughtrey	146.5	-1.66	6th
Diana Murphy	146.0	-1.68	8th
Thomas Nelson	140.1	-1.94*	9th
Daniel Manion	122.0	-2.71**	7th

** Indicates a Z-score of 1.96 or higher (representing a two-sided probability of <5% for a normal distribution).

* Indicates a Z-score of 1.645 or higher (representing a two-sided probability of <10% for a normal distribution).

The number of published opinions for each judge is adjusted so that the mean number of total opinions for each circuit is equal to 185.2 (the unadjusted mean number of total opinions for the seventh Circuit).

Summary Statistics for (A) (n=98): Mean = 185.2; Median = 182.05; Standard Deviation = 23.296; Kurtosis = 1.263; Skewness = 0.543.

Chi-squared test of the null hypothesis that the distribution of circuits is identical between the top judges and the bottom judges for the number of published opinions adjusted for intercircuit differences (A): $\chi^2 = 5.253$ (11 d.f.) ($p \leq 0.918$). Top judges are defined as those who are in the top fifty percent of judges based on the number of published opinions (majority, concurrences, and dissents) adjusted for intercircuit differences (A). Bottom judges are defined as those who are in the bottom fifty percent of judges based on the number of published opinions (majority, concurrences, and dissents) adjusted for intercircuit differences (A).

TABLE D. Citations to opinions published, 1998–2000

	(A) Total Outside Circuit	Z-Score of Normalized (A)	(B) SCT	Z-Score of Normalized (B)	(C) Law Review and Periodical	Z-Score of Normalized (C)	(D) Self-Citations	Z-Score of Normalized (D)	(E) Circuit
Richard Posner	1406	2.61**	16	2.31**	1033	2.41**	392	2.35**	7th
Frank Easterbrook	1340	2.52**	14	2.11**	790	1.83*	257	1.95*	7th
Sandra Lynch	1023	1.99**	5	0.62	684	1.52	178	1.60	1st
Bruce Selya	949	1.85*	3	-0.04	727	1.65*	364	2.28**	1st
Paul Kelly	799	1.51	0	-2.29**	388	0.30	103	1.07	10th
Michael Kanne	768	1.44	4	0.32	512	0.90	181	1.61	7th
Joel Flaum	743	1.37	3	-0.04	613	1.29	126	1.27	7th
Kenneth Ripple	730	1.34	4	0.32	545	1.03	168	1.54	7th
Diane Wood	678	1.20	3	-0.04	513	0.90	127	1.27	7th
J. Harvie Wilkinson	662	1.15	4	0.32	648	1.41	23	-0.36	4th
Edward Earl Carnes	648	1.11	4	0.32	448	0.61	113	1.16	11th
Jerry Smith	637	1.08	3	-0.04	622	1.32	53	0.44	5th
David Ebel	634	1.07	4	0.32	503	0.86	123	1.24	10th
Paul Niemeyer	623	1.03	10	1.60	390	0.31	46	0.31	4th
John Walker	616	1.01	4	0.32	667	1.47	36	0.07	2d
Stephen Reinhardt	605	0.98	11	1.75*	788	1.83*	66	0.65	9th
Pasco Bowman	573	0.87	4	0.32	487	0.79	42	0.22	8th
Carolyn Dineen King	572	0.87	5	0.62	490	0.80	60	0.56	5th
Guido Calabresi	566	0.85	7	1.09	604	1.25	55	0.48	2d

	(A) Total Outside Circuit	Z-Score of Normalized (A)	(B) SCT	Z-Score of Normalized (B)	(C) Law Review and Periodical	Z-Score of Normalized (C)	(D) Self-Citations	Z-Score of Normalized (D)	(E) Circuit
E. Grady Jolly	554	0.80	2	-0.51	422	0.48	20	-0.49	5th
Karen Williams	552	0.80	8	1.28	437	0.56	22	-0.40	4th
Juan Torruella	544	0.77	3	-0.04	444	0.59	84	0.88	1st
Roger Wollman	541	0.76	5	0.62	414	0.44	246	1.90	8th
Frank Hull	525	0.70	6	0.87	241	-0.73	21	-0.44	11th
Anthony Scirica	523	0.69	2	-0.51	236	-0.77	19	-0.54	3d
James Loken	518	0.67	3	-0.04	446	0.60	61	0.57	8th
Karen Nelson Moore	517	0.67	4	0.32	403	0.38	160	1.49	6th
Danny Boggs	507	0.63	6	0.87	424	0.49	39	0.15	6th
Gerald Bard Tjoflat	507	0.63	3	-0.04	498	0.84	65	0.63	11th
Dennis Jacobs	503	0.62	4	0.32	450	0.62	36	0.07	2d
Fortunato Benavides	469	0.48	2	-0.51	456	0.65	22	-0.40	5th
Daniel Manion	467	0.47	4	0.32	386	0.29	80	0.83	7th
Michael Murphy	454	0.42	6	0.87	368	0.19	109	1.13	10th
John Coffey	446	0.38	2	-0.51	465	0.69	173	1.57	7th
Michael Boudin	443	0.37	6	0.87	346	0.05	82	0.86	1st
Deanell Reece Tacha	442	0.37	2	-0.51	509	0.89	83	0.87	10th
Diana Gribbon Motz	430	0.31	4	0.32	253	-0.62	18	-0.59	4th
Eric Clay	418	0.26	3	-0.04	383	0.27	113	1.16	6th
Stanley Birch	417	0.25	7	1.09	401	0.37	25	-0.28	11th
Harold DeMoss	412	0.23	6	0.87	293	-0.30	34	0.02	5th

	(A) Total Outside Circuit	Z-Score of Normalized (A)	(B) SCT	Z-Score of Normalized (B)	(C) Law Review and Periodical	Z-Score of Normalized (C)	(D) Self-Citations	Z-Score of Normalized (D)	(E) Circuit
Ronald Lee Gilman	409	0.22	4	0.32	400	0.37	75	0.77	6th
Emilio Garza	406	0.20	6	0.87	385	0.28	22	-0.40	5th
Terrence Evans	405	0.20	4	0.32	413	0.44	31	-0.07	7th
Dolores Sloviter	403	0.19	5	0.62	394	0.33	27	-0.20	3d
Jose Cabranes	396	0.15	1	-1.16	354	0.10	16	-0.70	2d
Richard Nygaard	394	0.14	1	-1.16	191	-1.23	11	-1.06	3d
Diana Murphy	393	0.14	0	-2.29**	386	0.29	26	-0.24	8th
Ilana Rovner	392	0.13	3	-0.04	376	0.23	66	0.65	7th
Diarmuid O'Scannlain	386	0.10	6	0.87	741	1.69	29	-0.13	9th
Edith Jones	380	0.07	3	-0.04	589	1.20	25	-0.28	5th
J. Michael Luttig	378	0.06	8	1.28	476	0.74	16	-0.70	4th
Morris Arnold	369	0.02	6	0.87	357	0.12	60	0.56	8th
Rhesa Barksdale	368	0.01	5	0.62	334	-0.02	26	-0.24	5th
A. Wallace Tashima	367	0.01	4	0.32	317	-0.13	34	0.02	9th
Patrick Higginbotham	364	-0.01	5	0.62	544	1.03	29	-0.13	5th
Jacques Wiener	349	-0.09	3	-0.04	401	0.37	21	-0.44	5th
Carlos Lucero	346	-0.11	6	0.87	305	-0.22	72	0.73	10th
David Tatel	345	-0.11	7	1.09	212	-1.00	40	0.17	D.C.
William Wilkins	341	-0.14	5	0.62	374	0.22	29	-0.13	4th
Stephanie Seymour	340	-0.14	5	0.62	306	-0.21	43	0.24	10th
Carl Stewart	319	-0.27	4	0.32	527	0.96	36	0.07	5th

	(A) Total Outside Circuit	Z-Score of Normalized (A)	(B) SCT	Z-Score of Normalized (B)	(C) Law Review and Periodical	Z-Score of Normalized (C)	(D) Self-Citations	Z-Score of Normalized (D)	(E) Circuit
James Edmondson	317	-0.28	7	1.09	390	0.31	16	-0.70	11th
W. Eugene Davis	309	-0.33	2	-0.51	440	0.57	9	-1.25	5th
Sidney Thomas	306	-0.35	6	0.87	314	-0.16	34	0.02	9th
R. Guy Cole	293	-0.43	6	0.87	212	-1.00	56	0.49	6th
Mary Beck Briscoe	284	-0.49	6	0.87	247	-0.67	30	-0.10	10th
Alice Batchelder	281	-0.51	4	0.32	201	-1.12	16	-0.70	6th
Stephen Trott	281	-0.51	4	0.32	422	0.48	26	-0.24	9th
Harry Pregerson	278	-0.53	2	-0.51	241	-0.73	23	-0.36	9th
Rosemary Barkett	276	-0.55	4	0.32	251	-0.64	15	-0.76	11th
Robert Henry	276	-0.55	4	0.32	232	-0.81	45	0.28	10th
Samuel Alito	263	-0.64	4	0.32	240	-0.73	12	-0.98	3d
Fred Parker	261	-0.66	3	-0.04	285	-0.36	14	-0.83	2d
M. Blane Michael	248	-0.76	3	-0.04	200	-1.13	17	-0.64	4th
Robert Anderson	246	-0.77	1	-1.16	273	-0.46	11	-1.06	11th
Theodore McKee	244	-0.79	1	-1.16	172	-1.45	22	-0.40	3d
Marjorie Rendell	244	-0.79	0	-2.29**	215	-0.97	17	-0.64	3d
Stanley Marcus	237	-0.84	1	-1.16	334	-0.02	51	0.40	11th
Jane Roth	232	-0.89	2	-0.51	199	-1.14	13	-0.90	3d
Mary Schroeder	230	-0.90	1	-1.16	237	-0.76	3	-2.30**	9th
Martha Daughtrey	224	-0.95	2	-0.51	160	-1.61	2	-2.68**	6th

	(A) Total Outside Circuit	Z-Score of Normalized (A)	(B) SCT	Z-Score of Normalized (B)	(C) Law Review and Periodical	Z-Score of Normalized (C)	(D) Self-Citations	Z-Score of Normalized (D)	(E) Circuit
Susan Harrell Black	220	-0.99	1	-1.16	185	-1.30	8	-1.36	11th
Arthur Randolph	216	-1.02	2	-0.51	292	-0.31	36	0.07	D.C.
Joel Fredrick Dubina	214	-1.04	4	0.32	152	-1.72*	8	-1.36	11th
Rogers, Judith Rogers	214	-1.04	2	-0.51	193	-1.20	24	-0.31	D.C.
Michael Daly Hawkins	214	-1.04	4	0.32	196	-1.17	11	-1.06	9th
James Dennis	202	-1.15	0	-2.29**	204	-1.08	18	-0.59	5th
Douglas Ginsburg	193	-1.24	2	-0.51	263	-0.54	32	-0.04	D.C.
Harry Edwards	177	-1.41	1	-1.16	229	-0.84	22	-0.40	D.C.
Merrick Garland	169	-1.50	1	-1.16	176	-1.40	72	0.73	D.C.
H. Emory Widener	169	-1.50	1	-1.16	305	-0.22	4	-2.02**	4th
Thomas Nelson	167	-1.52	2	-0.51	117	-2.28	2	-2.68**	9th
Andrew Kleinfeld	157	-1.64	0	-2.29**	164	-1.56	19	-0.54	9th
Alex Kozinski	149	-1.74*	4	0.32	216	-0.96	10	-1.15	9th
David Bryan Sentelle	148	-1.76*	2	-0.51	148	-1.78*	26	-0.24	D.C.
Boyce Martin	124	-2.10**	0	-2.29**	138	-1.93*	17	-0.64	6th
Pamela Ann Rymer	112	-2.30**	0	-2.29**	206	-1.06	3	-2.30**	9th
Karen Henderson	109	-2.35**	0	-2.29**	116	-2.30**	35	0.04	D.C.

** Indicates a Z-score of 1.96 or higher (representing a two-sided probability of <5% for a normal distribution).

* Indicates a Z-score of 1.645 or higher (representing a two-sided probability of <10% for a normal distribution).

Outside-circuit citations are measured as of May 31, 2003. Normalized (A) is equal to LN(Total Outside-Circuit Citations). Normalized (B) is equal to LN(1+SCT Citations). Normalized (C) is equal to LN(Law Review and Periodical Citations). Normalized (D) is equal to LN(Self Citations).

Summary statistics for (A) (n=98): Mean = 417.3; Median = 383.0; Standard Deviation = 229.5; Kurtosis = 5.028; Skewness = 1.795.

Summary statistics for normalized (A) (n=98): Mean = 5.903; Median = 5.948; Standard Deviation = 0.515; Kurtosis = 0.025; Skewness = -0.020.

Summary statistics for (B) (n=98): Mean = 3.837; Median = 4.000; Standard Deviation = 2.757; Kurtosis = 4.583; Skewness = 1.547.

Summary statistics for normalized (B) (n=98): Mean = 1.410; Median = 1.609; Standard Deviation = 0.616; Kurtosis = 0.438; Skewness = -0.657.

Summary statistics for (C) (n=98): Mean = 374.2; Median = 375.0; Standard Deviation = 172.0; Kurtosis = 1.408; Skewness = 0.992.

Summary statistics for normalized (C) (n=98): Mean = 5.822; Median = 5.927; Standard Deviation = 0.464; Kurtosis = -0.497; Skewness = -0.148.

Summary statistics for (D) (n=98): Mean = 56.51; Median = 30.50; Standard Deviation = 69.05; Kurtosis = 9.287; Skewness = 2.807.

Summary statistics for normalized (D) (n=98): Mean = 3.508; Median = 3.418; Standard Deviation = 1.049; Kurtosis = 0.509; Skewness = -0.149.

Chi-squared test of the null hypothesis that the distribution of circuits is identical between the top judges and the bottom judges for the number of outside-circuit citations (A): $\chi^2 = 31.553$ (11 d.f.) ($p \leq 0.001$). Top judges are defined as those who are in the top fifty percent of judges based on the number of outside-circuit citations (A). Bottom judges are defined as those who are in the bottom fifty percent of judges based on the number of outside-circuit citations (A).

TABLE E. Outside-circuit citations to opinions controlling for total number of opinions published, 1998–2000

	(A) Outside-Circuit Citations to Judge's Top Twenty Opinions	Z-Score of Normalized (A)	(B) Average Outside-Circuit Citations per Majority Opinion	Z-Score of Normalized (B)	(C) Circuit
Sandra Lynch	734	2.56**	9.03	1.73*	1st
Frank Easterbrook	667	2.33**	6.25	0.73	7th
Paul Kelly	654	2.28**	9.85	1.97**	10th
Richard Posner	570	1.95*	5.49	0.37	7th
Bruce Selya	516	1.71*	6.50	0.83	1st
Anthony Scirica	496	1.61	14.50	3.04**	3d
Frank Hull	460	1.43	10.90	2.25**	11th
Karen Williams	455	1.40	11.02	2.28**	4th
Edward Earl Carnes	444	1.34	8.92	1.70*	11th
J. Harvie Wilkinson	425	1.24	7.64	1.28	4th
John Walker	423	1.22	6.75	0.93	2d
E. Grady Jolly	415	1.18	7.65	1.28	5th
David Ebel	412	1.16	6.22	0.71	10th
Carolyn Dineen King	407	1.13	7.61	1.27	5th
Jerry Smith	389	1.02	5.30	0.27	5th
Stephen Reinhardt	374	0.93	6.33	0.76	9th
Guido Calabresi	371	0.91	6.67	0.90	2d
Pasco Bowman	371	0.91	4.16	-0.39	8th
Paul Niemeyer	365	0.87	6.77	0.94	4th
Dennis Jacobs	345	0.73	5.40	0.32	2d
Joel Flaum	342	0.71	3.86	-0.60	7th
Fortunato Benavides	341	0.70	5.64	0.44	5th
Kenneth Ripple	340	0.69	4.82	0.01	7th
Richard Nygaard	336	0.67	8.76	1.65*	3d
Michael Kanne	328	0.61	4.36	-0.27	7th
Diane Wood	327	0.60	3.87	-0.60	7th
J. Michael Luttig	322	0.56	6.05	0.63	4th
Danny Boggs	320	0.55	5.63	0.44	6th

	(A) Outside-Circuit Citations to Judge's Top Twenty Opinions	Z-Score of Normalized (A)	(B) Average Outside-Circuit Citations per Majority Opinion	Z-Score of Normalized (B)	(C) Circuit
Karen Nelson Moore	315	0.51	5.45	0.35	6th
Gerald Bard Tjoflat	313	0.49	5.11	0.17	11th
Rhesa Barksdale	309	0.46	6.13	0.67	5th
Eric Clay	309	0.46	6.65	0.89	6th
Diana Gribbon Motz	307	0.45	6.87	0.98	4th
William Wilkins	295	0.35	7.19	1.11	4th
Stanley Birch	292	0.33	5.05	0.14	11th
Harold DeMoss	288	0.29	5.24	0.24	5th
Diarmuid O'Scannlain	284	0.26	4.07	-0.45	9th
Juan Torruella	281	0.23	4.20	-0.37	1st
Dolores Sloviter	277	0.20	5.04	0.13	3d
Deanell Reece Tacha	277	0.20	4.79	-0.01	10th
Ronald Lee Gilman	274	0.17	4.05	-0.47	6th
Roger Wollman	274	0.17	3.51	-0.86	8th
Emilio Garza	272	0.15	5.35	0.30	5th
Jacques Loeb Wiener	267	0.11	4.64	-0.09	5th
Daniel Manion	265	0.09	4.56	-0.14	7th
David Tatel	265	0.09	5.11	0.17	D.C.
Edith Jones	262	0.06	4.74	-0.04	5th
James Loken	258	0.03	3.51	-0.86	8th
A. Wallace Tashima	257	0.02	4.79	-0.01	9th
Stephanie Seymour	256	0.01	5.37	0.30	10th
James Larry Edmondson	255	0.00	6.22	0.71	11th
Michael Murphy	253	-0.02	4.45	-0.21	10th
W. Eugene Davis	250	-0.05	4.46	-0.20	5th
Alice Batchelder	250	-0.05	5.83	0.53	6th
Diana Murphy	245	-0.10	3.69	-0.73	8th
Carlos Lucero	240	-0.15	4.89	0.05	10th
Sidney Thomas	233	-0.22	4.51	-0.17	9th
Patrick Higginbotham	232	-0.23	3.81	-0.64	5th
Ilana Rovner	232	-0.23	3.18	-1.13	7th
Terrence Evans	230	-0.25	3.00	-1.29	7th
Jose Cabranes	229	-0.26	5.58	0.41	2d

	(A) Outside-Circuit Citations to Judge's Top Twenty Opinions	Z-Score of Normalized (A)	(B) Average Outside-Circuit Citations per Majority Opinion	Z-Score of Normalized (B)	(C) Circuit
Michael Boudin	223	-0.33	3.31	-1.03	1st
John Coffey	222	-0.34	2.75	-1.54	7th
Theodore McKee	218	-0.38	5.90	0.57	3d
R. Guy Cole	218	-0.38	5.29	0.27	6th
Robert Anderson III	218	-0.38	5.55	0.40	11th
M. Blane Michael	217	-0.39	6.45	0.81	4th
Martha Craig Daughtrey	212	-0.45	5.89	0.56	6th
Mary Beck Briscoe	207	-0.51	4.76	-0.02	10th
Samuel Alito	205	-0.53	4.94	0.08	3d
Rosemary Barkett	205	-0.53	4.17	-0.39	11th
Robert Henry	201	-0.58	4.91	0.06	10th
Stephen Trott	198	-0.62	3.23	-1.09	9th
Marjorie Rendell	196	-0.64	6.21	0.70	3d
Jane Roth	196	-0.64	5.54	0.39	3d
Fred Parker	193	-0.68	4.50	-0.18	2d
Stanley Marcus	190	-0.72	4.84	0.02	11th
Morris Arnold	188	-0.74	2.36	-1.95*	8th
Mary Schroeder	187	-0.76	3.38	-0.96	9th
Joel Fredrick Dubina	187	-0.76	5.35	0.30	11th
Susan Harrell Black	172	-0.96	4.47	-0.20	11th
Judith Rogers	167	-1.03	2.86	-1.42	D.C.
Harry Pregerson	166	-1.04	3.39	-0.96	9th
Arthur Randolph	164	-1.07	3.48	-0.89	D.C.
Carl Stewart	163	-1.09	3.13	-1.18	5th
Michael Daly Hawkins	163	-1.09	3.55	-0.83	9th
James Dennis	156	-1.20	3.78	-0.66	5th
Douglas Ginsburg	149	-1.31	2.49	-1.80*	D.C.
Thomas Nelson	148	-1.32	4.18	-0.38	9th
Harry Edwards	145	-1.37	2.72	-1.56	D.C.
Merrick Garland	143	-1.41	3.25	-1.07	D.C.
H. Emory Widener	139	-1.48	3.29	-1.04	4th
Alex Kozinski	137	-1.51	3.65	-0.75	9th
Andrew Kleinfeld	133	-1.58	2.96	-1.33	9th

	(A) Outside-Circuit Citations to Judge's Top Twenty Opinions	Z-Score of Normalized (A)	(B) Average Outside-Circuit Citations per Majority Opinion	Z-Score of Normalized (B)	(C) Circuit
David Bryan Sentelle	113	-1.98**	2.45	-1.85*	D.C.
Boyce Martin	110	-2.04**	2.98	-1.32	6th
Karen Henderson	98	-2.32**	2.08	-2.30**	D.C.
Pamela Ann Rymer	96	-2.37**	2.47	-1.83*	9th

** Indicates a Z-score of 1.96 or higher (representing a two-sided probability of <5% for a normal distribution).

* Indicates a Z-score of 1.645 or higher (representing a two-sided probability of <10% for a normal distribution).

Outside-circuit citations are measured as of May 31, 2003. Normalized (A) is equal to LN(Outside-Circuit Citations to Judge's Top Twenty Opinions). Normalized B is equal to LN(Average Outside-Circuit Citations per Majority Opinion).

Summary statistics for (A) (n=98): Mean = 277.9; Median = 256.5; Standard Deviation = 121.2; Kurtosis = 2.608; Skewness = 1.382.

Summary statistics for normalized (A) (n=98): Mean = 5.543; Median = 5.547; Standard Deviation = 0.412; Kurtosis = 0.098; Skewness = 0.068.

Summary statistics for (B) (n=98): Mean = 5.137; Median = 4.861; Standard Deviation = 2.030; Kurtosis = 4.577; Skewness = 1.638.

Summary statistics for normalized (B) (n=98): Mean = 1.569; Median = 1.581; Standard Deviation = 0.364; Kurtosis = 0.291; Skewness = 0.216.

Chi-squared test of the null hypothesis that the distribution of circuits is identical between the top judges and the bottom judges for the number of outside-circuit citations to the top twenty citation-receiving cases (A): $\chi^2 = 15.466$ (11 d.f.) ($p \leq 0.169$). Top judges are defined as those who are in the top fifty percent of judges based on the number of outside-circuit citations to the top twenty citation-receiving cases (A). Bottom judges are defined as those who are in the bottom fifty percent of judges based on the number of outside-circuit citations to the top twenty citation-receiving cases (A).

TABLE F. Invocations to opinions published, 1998–2000

	(A) Total Invocations	Z-Score of Normalized (A)	(B) Average Invocations per Opinion	Z-Score of Normalized (B)	(C) Percent of Invocations Attributable to Majority Opinions	(D) Circuit
Richard Posner	176	3.90**	0.65	6.68**	97.7%	7th
Frank Easterbrook	103	3.36**	0.44	4.67**	99.0%	7th
Guido Calabresi	23	1.85*	0.23	2.35**	91.3%	2d
J. Harvie Wilkinson	19	1.66*	0.18	1.73*	73.7%	4th
Michael Boudin	13	1.30	0.10	0.70	84.6%	1st
Patrick Higginbotham	12	1.22	0.12	0.96	41.7%	5th
Diarmuid O'Scannlain	11	1.14	0.08	0.43	54.5%	9th
Edith Jones	11	1.14	0.11	0.83	81.8%	5th
Diane Wood	10	1.05	0.05	0.01	20.0%	7th
J. Michael Luttig	10	1.05	0.12	0.96	80.0%	4th
Kenneth Ripple	9	0.95	0.05	0.01	55.6%	7th
John Walker	9	0.95	0.09	0.56	77.8%	2d
Gerald Bard Tjoflat	9	0.95	0.08	0.43	100.0%	11th
Morris Arnold	8	0.84	0.05	0.01	12.5%	8th
Bruce Selya	8	0.84	0.05	0.01	75.0%	1st
James Larry Edmondson	8	0.84	0.14	1.23	100.0%	11th
Ilana Rovner	8	0.84	0.05	0.01	12.5%	7th
Ronald Lee Gilman	8	0.84	0.06	0.15	25.0%	6th
Sandra Lynch	7	0.72	0.06	0.15	71.4%	1st
Carolyn Dineen King	7	0.72	0.08	0.43	71.4%	5th
Andrew Kleinfeld	7	0.72	0.09	0.56	42.9%	9th
Paul Niemeyer	7	0.72	0.06	0.15	71.4%	4th
Stephen Reinhardt	7	0.72	0.05	0.01	0.0%	9th
Eric Clay	7	0.72	0.09	0.56	71.4%	6th
Jose Cabranes	7	0.72	0.09	0.56	71.4%	2d
Terrence Evans	6	0.58	0.04	-0.13	100.0%	7th
Karen Nelson Moore	6	0.58	0.05	0.01	16.7%	6th

	(A) Total Invocations	Z-Score of Normalized (A)	(B) Average Invocations per Opinion	Z-Score of Normalized (B)	(C) Percent of Invocations Attributable to Majority Opinions	(D) Circuit
Samuel Alito	5	0.42	0.07	0.29	40.0%	3d
Daniel Manion	5	0.42	0.04	-0.13	40.0%	7th
Stephen Trott	5	0.42	0.05	0.01	60.0%	9th
Danny Boggs	5	0.42	0.04	-0.13	60.0%	6th
Alex Kozinski	5	0.42	0.07	0.29	100.0%	9th
Dennis Jacobs	5	0.42	0.05	0.01	0.0%	2d
Arthur Raymond Randolph	5	0.42	0.06	0.15	60.0%	D.C.
Judith Ann Wilson Rogers	5	0.42	0.06	0.15	80.0%	D.C.
Fortunato Pedro Benavides	5	0.42	0.05	0.01	60.0%	5th
Emilio Garza	5	0.42	0.05	0.01	0.0%	5th
David Tatel	5	0.42	0.06	0.15	60.0%	D.C.
James Loken	5	0.42	0.03	-0.27	20.0%	8th
Sidney Thomas	4	0.24	0.05	0.01	25.0%	9th
Jane Roth	4	0.24	0.07	0.29	25.0%	3d
Stephanie Seymour	4	0.24	0.06	0.15	75.0%	10th
James Dennis	4	0.24	0.04	-0.13	25.0%	5th
Diana Gribbon Motz	4	0.24	0.05	0.01	50.0%	4th
R. Guy Cole	4	0.24	0.06	0.15	100.0%	6th
Stanley Birch	4	0.24	0.04	-0.13	50.0%	11th
Edward Earl Carnes	3	0.01	0.03	-0.27	100.0%	11th
Joel Flaum	3	0.01	0.01	-0.56	66.7%	7th
Marjorie Rendell	3	0.01	0.05	0.01	33.3%	3d
Rosemary Barkett	3	0.01	0.03	-0.27	66.7%	11th
Jerry Smith	3	0.01	0.02	-0.41	0.0%	5th
Thomas Nelson	3	0.01	0.07	0.29	0.0%	9th
Harry Pregerson	3	0.01	0.03	-0.27	66.7%	9th
M. Blane Michael	3	0.01	0.05	0.01	66.7%	4th
Michael Kanne	3	0.01	0.02	-0.41	66.7%	7th
Anthony Scirica	2	-0.29	0.05	0.01	100.0%	3d
Mary Schroeder	2	-0.29	0.03	-0.27	50.0%	9th
Theodore McKee	2	-0.29	0.04	-0.13	0.0%	3d

	(A) Total Invocations	Z-Score of Normalized (A)	(B) Average Invocations per Opinion	Z-Score of Normalized (B)	(C) Percent of Invocations Attributable to Majority Opinions	(D) Circuit
William Wilkins	2	-0.29	0.03	-0.27	50.0%	4th
A. Wallace Tashima	2	-0.29	0.02	-0.41	0.0%	9th
John Coffey	2	-0.29	0.01	-0.56	50.0%	7th
Harry Edwards	2	-0.29	0.03	-0.27	50.0%	D.C.
Jacques Loeb Wiener	2	-0.29	0.02	-0.41	0.0%	5th
Susan Harrell Black	2	-0.29	0.03	-0.27	0.0%	11th
H. Emory Widener	2	-0.29	0.03	-0.27	0.0%	4th
Karen Williams	2	-0.29	0.03	-0.27	50.0%	4th
Frank Hull	2	-0.29	0.04	-0.13	50.0%	11th
Carlos Lucero	2	-0.29	0.02	-0.41	0.0%	10th
David Ebel	1	-0.71	0.01	-0.56	100.0%	10th
E. Grady Jolly	1	-0.71	0.01	-0.56	0.0%	5th
Juan Torruella	1	-0.71	0.01	-0.56	0.0%	1st
Harold DeMoss	1	-0.71	0.01	-0.56	0.0%	5th
Douglas Ginsburg	1	-0.71	0.01	-0.56	0.0%	D.C.
Dolores Sloviter	1	-0.71	0.01	-0.56	100.0%	3d
David Bryan Sentelle	1	-0.71	0.01	-0.56	100.0%	D.C.
Pamela Ann Rymer	1	-0.71	0.01	-0.56	0.0%	9th
Mary Beck Briscoe	1	-0.71	0.01	-0.56	0.0%	10th
Alice Batchelder	1	-0.71	0.01	-0.56	100.0%	6th
Pasco Bowman	1	-0.71	0.01	-0.56	100.0%	8th
Roger Wollman	1	-0.71	0.01	-0.56	100.0%	8th
Robert Lanier Anderson	0	-1.42	0.00	-0.71	–	11th
Carl Stewart	0	-1.42	0.00	-0.71	–	5th
Merrick Garland	0	-1.42	0.00	-0.71	–	D.C.
Richard Nygaard	0	-1.42	0.00	-0.71	–	3d
W. Eugene Davis	0	-1.42	0.00	-0.71	–	5th
Joel Fredrick Dubina	0	-1.42	0.00	-0.71	–	11th
Paul Kelly	0	-1.42	0.00	-0.71	–	10th
Fred Parker	0	-1.42	0.00	-0.71	–	2d
Michael Murphy	0	-1.42	0.00	-0.71	–	10th

	(A) Total Invocations	Z-Score of Normalized (A)	(B) Average Invocations per Opinion	Z-Score of Normalized (B)	(C) Percent of Invocations Attributable to Majority Opinions	(D) Circuit
Martha Craig Daughtrey	0	-1.42	0.00	-0.71	–	6th
Robert Henry	0	-1.42	0.00	-0.71	–	10th
Karen LeCraft Henderson	0	-1.42	0.00	-0.71	–	D.C.
Diana Murphy	0	-1.42	0.00	-0.71	–	8th
Michael Daly Hawkins	0	-1.42	0.00	-0.71	–	9th
Boyce Martin	0	-1.42	0.00	-0.71	–	6th
Rhesa Hawkins Barksdale	0	-1.42	0.00	-0.71	–	5th
Deanell Reece Tacha	0	-1.42	0.00	-0.71	–	10th
Stanley Marcus	0	-1.42	0.00	-0.71	–	11th

** Indicates a Z-score of 1.96 or higher (representing a two-sided probability of <5% for a normal distribution).

* Indicates a Z-score of 1.645 or higher (representing a two-sided probability of <10% for a normal distribution).

Normalized (A) is equal to LN(1+Invocations). Normalized (B) is equal to LN(1+Average Invocations per Opinion).

Summary statistics for (A) (n=98): Mean = 6.827; Median = 3.000; Standard Deviation = 20.36; Kurtosis = 54.685; Skewness = 7.145.

Summary statistics for normalized (A) (n=98): Mean = 1.379; Median = 1.386; Standard Deviation = 0.973; Kurtosis = 2.091; Skewness = 0.750.

Summary statistics for (B) (n=98): Mean = 0.053; Median = 0.038; Standard Deviation = 0.084; Kurtosis = 31.800; Skewness = 5.085.

Summary statistics for normalized (B) (n=98): Mean = 0.049; Median = 0.037; Standard Deviation = 0.068; Kurtosis = 24.700; Skewness = 4.372.

Chi-squared test of the null hypothesis that the distribution of circuits is identical between the top judges and the bottom judges for the number of invocations (A): $\chi^2 = 13.863$ (11 d.f.) ($p \leq 0.241$). Top judges are defined as those who are in the top fifty percent of judges based on the number of invocations (A). Bottom judges are defined as those who are in the bottom fifty percent of judges based on the number of invocations (A).

TABLE G. Number of opposing opinions and independence ratings for the period, 1998–2000

	(A) Independence Rating	(B) Number of Dissents and Concurrences	(C) Adjusted Dissents and Concurrences for Intercircuit Differences	Z-Score for (C)	(D) Circuit
David Ebel	0.000	12	24.5	0.16	10th
Robert Anderson	0.000	10	23.7	0.06	11th
J. Michael Luttig	-0.006	19	26.2	0.35	4th
Samuel Alito	-0.015	18	28.7	0.65	3d
Carl Stewart	-0.017	2	10.4	-1.49	5th
E. Grady Jolly	-0.018	10	18.4	-0.56	5th
Juan Torruella	-0.018	10	26.9	0.44	1st
Diane Wood	-0.018	21	27.0	0.44	7th
Richard Posner	0.019	15	21.0	-0.26	7th
Edward Earl Carnes	-0.022	14	27.7	0.53	11th
Anthony Scirica	-0.023	2	12.7	-1.22	3d
Mary Schroeder	0.023	5	5.0	-2.12**	9th
Terrence Evans	0.024	18	24.0	0.09	7th
Harold DeMoss	0.024	30	38.4	1.78	5th
Daniel Manion	0.028	20	26.0	0.33	7th
Merrick Garland	-0.037	3	14.5	-1.01	D.C.
Frank Easterbrook	-0.042	20	26.0	0.33	7th
Sandra Lynch	-0.043	7	23.9	0.09	1st
Stephen Trott	-0.044	19	19.0	-0.49	9th
Joel Flaum	-0.044	10	16.0	-0.84	7th
Edith Jones	-0.045	21	29.4	0.73	5th
Marjorie Rendell	-0.049	16	26.7	0.42	3d
Richard Nygaard	0.049	12	22.7	-0.05	3d
Douglas Ginsburg	0.052	2	13.5	-1.12	D.C.
Sidney Thomas	-0.053	18	18.0	-0.60	9th

	(A) Independence Rating	(B) Number of Dissents and Concurrences	(C) Adjusted Dissents and Concurrences for Intercircuit Differences	Z-Score for (C)	(D) Circuit
Rosemary Barkett	0.056	23	36.7	1.58	11th
Morris Arnold	-0.060	23	35.2	1.40	8th
Theodore McKee	-0.062	14	24.7	0.18	3d
William Wilkins	0.065	11	18.2	-0.58	4th
Carolyn Dineen King	-0.067	10	18.4	-0.56	5th
Michael Boudin	0.071	4	20.9	-0.26	1st
W. Eugene Davis	-0.071	1	9.4	-1.61	5th
Bruce Selya	-0.071	4	20.9	-0.26	1st
Dolores Sloviter	-0.077	12	22.7	-0.05	3d
Danny Boggs	-0.082	24	28.2	0.58	6th
Andrew Kleinfeld	0.087	30	30.0	0.80	9th
Jerry Smith	-0.093	14	22.4	-0.09	5th
James Edmondson	0.100	7	20.7	-0.29	11th
Joel Fredrick Dubina	0.100	4	17.7	-0.64	11th
J. Harvie Wilkinson	-0.102	17	24.2	0.12	4th
Paul Niemeyer	-0.102	21	28.2	0.58	4th
Patrick Higginbotham	-0.036	3	11.4	-1.38	5th
Kenneth Ripple	-0.107	31	37.0	1.61	7th
Jane Roth	0.112	13	23.7	0.07	3d
Ilana Rovner	0.112	30	36.0	1.50	7th
Paul Kelly	0.115	11	23.5	0.04	10th
A. Wallace Tashima	-0.122	23	23.0	-0.02	9th
Fred Parker	-0.124	8	20.0	-0.37	2d
Michael Murphy	-0.127	4	16.5	-0.77	10th

	(A) Independence Rating	(B) Number of Dissents and Concurrences	(C) Adjusted Dissents and Concurrences for Intercircuit Differences	Z-Score for (C)	(D) Circuit
Stephen Reinhardt	-0.128	48	48.0	2.90	9th
Alex Kozinski	0.135	27	27.0	0.45	9th
Thomas Nelson	0.135	6	6.0	-2.01*	9th
John Coffey	0.139	6	12.0	-1.31	7th
Stephanie Seymour	-0.150	5	17.5	-0.66	10th
James Dennis	-0.157	40	48.4	2.95**	5th
Diarmuid O'Scannlain	-0.165	45	45.0	2.55**	9th
David Bryan Sentelle	-0.166	17	28.5	0.63	D.C.
Harry Edwards	-0.170	7	18.5	-0.54	D.C.
Pamela Ann Rymer	-0.180	22	22.0	-0.14	9th
Martha Daughtrey	-0.181	8	12.2	-1.29	6th
Diana Gribbon Motz	-0.185	14	21.2	-0.23	4th
Robert Henry	-0.200	7	19.5	-0.42	10th
Dennis Jacobs	0.203	17	29.0	0.68	2d
Jacques Wiener	0.208	8	16.4	-0.79	5th
Arthur Randolph	0.208	16	27.5	0.51	D.C.
Mary Beck Briscoe	-0.210	26	38.5	1.80*	10th
Eric Clay	-0.212	19	23.2	0.00	6th
Karen Henderson	-0.219	25	36.5	1.56	D.C.
Judith Rogers	-0.227	6	17.5	-0.66	D.C.
Diana Murphy	0.228	5	17.2	-0.70	8th
John Walker	0.228	4	16.0	-0.84	2d
Ronald Lee Gilman	-0.229	25	29.2	0.70	6th
Harry Pregerson	-0.232	17	17.0	-0.72	9th

	(A) Independence Rating	(B) Number of Dissents and Concurrences	(C) Adjusted Dissents and Concurrences for Intercircuit Differences	Z-Score for (C)	(D) Circuit
Karen Nelson Moore	-0.233	36	40.2	1.99**	6th
Susan Harrell Black	-0.234	10	23.7	0.06	11th
Alice Batchelder	0.057	19	23.2	0.00	6th
Fortunato Benavides	-0.246	11	19.4	-0.44	5th
Emilio Garza	-0.254	31	39.4	1.90*	5th
H. Emory Widener	0.257	20	27.2	0.47	4th
Karen Williams	-0.259	9	16.2	-0.82	4th
Michael Daly Hawkins	-0.261	18	18.0	-0.60	9th
Boyce Martin	0.273	7	11.2	-1.40	6th
Rhesa Barksdale	-0.279	11	19.4	-0.44	5th
R. Guy Cole	-0.282	14	18.2	-0.58	6th
Pasco Bowman	-0.283	3	15.2	-0.93	8th
David Tatel	-0.287	17	28.5	0.63	D.C.
Frank Hull	0.292	6	19.7	-0.41	11th
Gerald Bard Tjoflat	-0.294	10	23.7	0.06	11th
Guido Calabresi	-0.307	17	29.0	0.68	2d
Jose Cabranes	-0.307	10	22.0	-0.14	2d
M. Blane Michael	-0.324	17	24.2	0.12	4th
Stanley Birch	0.333	10	23.7	0.06	11th
Deanell Reece Tacha	-0.389	7	19.5	-0.42	10th
Carlos Lucero	-0.400	13	25.5	0.28	10th
Michael Kanne	-0.417	1	7.0	-1.89*	7th
Roger Wollman	-0.449	4	16.2	-0.82	8th
James Loken	-0.496	20	32.2	1.05	8th
Stanley Marcus	0.542	1	14.7	-0.99	11th

** Indicates a Z-score of 1.96 or higher (representing a two-sided probability of <5% for a normal distribution).

* Indicates a Z-score of 1.645 or higher (representing a two-sided probability of <10% for a normal distribution).

Summary statistics for (A) (n=98): Mean = -0.062; Median = -0.057; standard deviation = 0.189; Kurtosis = 0.307; Skewness = 0.307.

Summary statistics for (B) (n=98): Mean = 14.469; Median = 13.00; standard deviation = 9.692; Kurtosis = 1.344; Skewness = 1.042.

Summary statistics for (C) (n=98): Mean = 23.170; Median = 22.869; standard deviation = 8.557; Kurtosis = 0.796; Skewness = 0.646.

Chi-squared test of the null hypothesis that the distribution of circuits is identical between the top judges and the bottom judges for the independence rating (A): $\chi^2 = 23.110$ (11 d.f.) ($p \leq 0.017$). Top judges are defined as those who are in the top fifty percent of judges based on the independence rating (A). Bottom judges are defined as those who are in the bottom fifty percent of judges based on the independence rating (A).

TABLE H. Composite ranking of judges using equal weighting of quality, productivity, and independence (includes only active circuit court judges sixty-five or younger in 2003)

	(A) Rank Based on Equal Weighted Composite Measure	(B) Z-Score Based on Equal Weighted Composite Measure	(C) Circuit
Richard Posner	1	3.77**	7th
Frank Easterbrook	2	2.93**	7th
J. Harvie Wilkinson	3	1.51	4th
Paul Niemeyer	4	1.34	4th
Jerry Smith	5	1.26	5th
David Ebel	6	1.21	10th
Edward Earl Carnes	7	1.20	11th
Diane Wood	8	1.14	7th
Danny Boggs	9	1.10	6th
J. Michael Luttig	10	1.08	4th
Sandra Lynch	11	1.02	1st
Karen Nelson Moore	12	0.89	6th
Edith Jones	13	0.88	5th
Morris Arnold	14	0.87	8th
Ronald Lee Gilman	15	0.77	6th
Samuel Alito	16	0.69	3d
Carolyn Dineen King	17	0.66	5th
Stephen Trott	18	0.66	9th
Rosemary Barkett	19	0.59	11th
Michael Boudin	20	0.57	1st
Kenneth Ripple	21	0.56	7th
Patrick Higginbotham	22	0.52	5th
Dennis Jacobs	23	0.50	2d
Paul Kelly	24	0.45	10th
Anthony Scirica	25	0.42	3d
John Walker	26	0.41	2d
Sidney Thomas	27	0.28	9th
James Edmondson	28	0.21	11th
Richard Nygaard	29	0.13	3d

	(A) Rank Based on Equal Weighted Composite Measure	(B) Z-Score Based on Equal Weighted Composite Measure	(C) Circuit
Marjorie Rendell	30	0.12	3d
Theodore McKee	31	0.06	3d
Diana Gribbon Motz	32	0.05	4th
Michael Murphy	33	0.04	10th
Eric Clay	34	0.03	6th
Carl Stewart	35	0.02	5th
Emilio Garza	36	0.01	5th
William Wilkins	37	-0.02	4th
Fortunato Benavides	38	-0.04	5th
Terrence Evans	39	-0.05	7th
Stanley Birch	40	-0.08	11th
Douglas Ginsburg	41	-0.08	D.C.
Mary Schroeder	42	-0.13	9th
Andrew Kleinfeld	43	-0.14	9th
David Tatel	44	-0.22	D.C.
Stephanie Seymour	45	-0.33	10th
Ilana Rovner	46	-0.33	7th
Arthur Randolph	47	-0.34	D.C.
Judith Rogers	48	-0.36	D.C.
Karen Williams	49	-0.45	4th
Frank Hull	50	-0.51	11th
Harry Edwards	51	-0.59	D.C.
Mary Beck Briscoe	52	-0.62	10th
Daniel Manion	53	-0.63	7th
Jose Cabranes	54	-0.64	2d
Alex Kozinski	55	-0.68	9th
Merrick Garland	56	-0.72	D.C.
David Bryan Sentelle	57	-0.74	D.C.
R. Guy Cole	58	-0.84	6th
James Loken	59	-0.85	8th
Susan Harrell Black	60	-0.85	11th
Joel Fredrick Dubina	61	-0.89	11th
Fred Parker	62	-0.95	2d
Michael Kanne	63	-0.99	7th

	(A) Rank Based on Equal Weighted Composite Measure	(B) Z-Score Based on Equal Weighted Composite Measure	(C) Circuit
Deanell Reece Tacha	64	-1.03	10th
Carlos Lucero	65	-1.17	10th
M. Blane Michael	66	-1.19	4th
Rhesa Barksdale	67	-1.19	5th
Robert Henry	68	-1.23	10th
Karen Henderson	69	-1.25	D.C.
Alice Batchelder	70	-1.25	6th
Michael Daly Hawkins	71	-1.34	9th
Martha Daughtrey	72	-1.38	6th
Pamela Ann Rymer	73	-1.40	9th
Stanley Marcus	74	-2.44**	11th

** Indicates a Z-score of 1.96 or higher (representing a two-sided probability of <5% for a normal distribution).

* Indicates a Z-score of 1.645 or higher (representing a two-sided probability of <10% for a normal distribution).

Equal weighted composite measure = 0.333Quality + 0.333Productivity – 0.333Independence.

Summary statistics for (A) (n=98): Mean = -9.832; Median = -10.373; Standard Deviation = 16.911; Kurtosis = 2.320; Skewness = 0.828.

TABLE I. Comparison of seniority of judges (ranked based on the outside-circuit citations to the top twenty citation-receiving opinions)

Panel A: Seniority of judges with highest top twenty citation counts

Judges with Highest Top Twenty Outside-Circuit Citation Count	**Seniority Quartile on the Circuit**	**Chief Judge Status in 1998–2000**	**Years Experience**
Sandra Lynch	4	0	3
Frank Easterbrook	2	0	13
Paul Kelly	2	0	6
Richard Posner	1	1	17
Bruce Selya	2	0	12
Anthony Scirica	1	0	11
Frank Hull	4	0	1
Karen Williams	3	0	6
Edward Earl Carnes	3	0	6
J. Harvie Wilkinson	1	1	14
Mean	**2.3**	**0.2**	**8.9**

Panel B: Seniority of eleven judges centered on the judge with the median number of outside-circuit citations for the top twenty opinions

Median Judges	**Seniority Quartile on the Circuit**	**Chief Judge Status in 1998–2000**	**Years Experience**
David Tatel	4	0	4
Edith Jones	2	0	13
James Loken	3	0	8
A. Wallace Tashima	4	0	2
Stephanie Seymour	1	1	19
James Larry Edmondson	2	0	12
Michael Murphy	4	0	3
Eugene Davis	2	0	15
Alice Batchelder	2	0	7
Diana Murphy	4	0	4
Carlos Lucero	4	0	3
Mean	**2.8**	**0.1**	**8.6**

Panel C: Seniority of judges with the lowest top twenty citation counts

Median Judges	Seniority Quartile on the Circuit	Chief Judge Status in 1998–2000	Years Experience
Thomas Nelson	3	0	8
Harry Edwards	1	1	18
Merrick Garland	4	0	1
H. Emory Widener	1	0	26
Alex Kozinski	2	0	13
Andrew Kleinfeld	3	0	7
David Bryan Sentelle	2	0	11
Boyce Martin	1	1	19
Karen Henderson	3	0	8
Pamela Ann Rymer	3	0	9
Mean	**2.3**	**0.2**	**12.0**

Seniority quartile 1 to 4 represents the top twenty-five percent to the bottom twenty-five percent quartiles respectively in the circuit based on seniority. Chief judge status = 1 if the judge was a chief judge of her respective circuit at any time during the 1998–2000 time period.

Unpaired t-test of difference between means of seniority quartile in Panels A and C: -1.21 (p<0.239).

Unpaired t-test of difference between means of seniority quartile in Panels A and B: 0.00 (p<1.000).

Unpaired t-test of difference between means of seniority quartile in Panels B and C: 1.27 (p<0.221).