Preliminary Draft 10-Nov-03
Please note that the numbers are preliminary and subject to change.

Who Would Win a Tournament of Judges?

Stephen Choi[*]
Mitu Gulati [**]

[*] 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, Édeanna Johnson-Chebbi, Margaret Rodgers, Rishi Sharma, Jennifer Dukart, and Alice Kuo provided research assistance. Kimberly Brickell deserves special thanks for her work. Aspects of this draft benefited from discussions with Scott Baker, Lee Epstein, Tracey George, Prea Gulati, Vicki Jackson, Mike Klarman, Kim Krawiec, Kaleb Michaud, Un Kyung Park, Jim Rossi, Michael Solimine, and Eric Talley. For comments on the draft, we are grateful to Suzette Baker, James Brudney, Michael Gerhardt, Steve Goldberg, Albert Lauber, Bill Marshall, Gerry Spann, David Vladeck, and the student participants in the Interdisciplinary Workshop at the Georgetown Law Center.

Disclosure: Funding for this project was provided entirely by our respective law schools. In addition, one of us was a law clerk to two of the judges in the sample: Samuel Alito of the 3rd Circuit and Sandra Lynch of the 1st Circuit. The individuals and institutions mentioned above bear no responsibility for the ideas expressed here.

Abstract

The impetus for the Article was frustration with the current judicial appointments process. Both sides claim that their candidates are the "most meritorious" and yet there is seldom any real discussion regarding merit (or even what should constitute merit). Instead, the discussion moves almost immediately to the candidates' likely positions on a set of hot button political issues like abortion, gun control, and the death penalty. One side (these days, the Republicans) claims that it is proposing certain candidates based on merit, while the other (the Democrats) claims that the real reason for pushing those candidates is their ideology and, in particular, their likely votes on certain 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 the 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 those claims helps make transparent when 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. Indeed, 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.

The 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 plausible measures of merit focusing on (a) productivity, (b) opinion quality, and (c) judicial independence. While not perfect, our measures do not have to meet this standard to prove useful. They only need to interject a greater focus on merit in the current nomination process (thereby flushing out previously non-transparent 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 or the other of his favorite circuit court judges to be his nominee for the Supreme Court.

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] As a result, 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 probably unfold: First, 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. Focus will be on hot button political issues such as abortion, the death penalty, and affirmative action, and what the candidate's expected position will be on each. Discussion of the candidate's broader qualifications will be pushed to the background. At best, the press accounts will carry brief mention of the candidate's employment history and the name of the law school she 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, including that which the President touted at the outset: *her career as a circuit court judge*.[2]

The genesis of this project lies in our frustration with the current appointments process. As best we can tell, the entire focus in analyzing a candidate's qualifications is in predicting her 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] With a number of recent high-profile nominations to the

---

[1] *E.g.,* Sheryl Gay Stolberg, The War Over Abortion Moves to a Smaller Stage, N.Y. TIMES, Section 4, Page 4, October 26, 2003 ("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, and Robert A. Carp, George W. Bush's Potential Supreme Court Nominees: What Impact Might They Have? 85 JUDICATURE 278 (2002).

[2] In recent years, a norm appears to have developed where 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 et al., 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, WASHINGTON POST,

federal courts of appeals—Miguel Estrada, Priscilla Owen, 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 it is their opponents who are the ideologues because they are blocking a highly qualified candidate.[4]

For purposes of this Article, we make two observations regarding the above scenario. First, both sides seem to perceive merit and not political ideology as what their *constituents* want selections to be based upon. 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 meritorious. Second, there is little quarrel over the merit of the nominees—we suspect, because the politicians care little about it and because the public has no means of evaluating it.[5] Instead, claims of merit are largely ignored and the quarrel takes place over ideology, 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?

Wednesday, July 30, 2003, 2003 WL 56509262 (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"); Statement on the Senate Filibuster of Judicial Nominees by George W. Bush, Weekly Compilation of Presidential Documents, 2003 WL 1397 3558 (claiming that his "highly qualified nominees" with "stellar records" are being blocked without justification); Chocola Supports Bush Court Nominee, SOUTH BEND TRIBUNE, Friday February 14, 2003, 2003 WL 9896880, (asserting that there is "no question that Miguel Estrada is highly qualified to serve on the federal bench," but providing little evidence beyond Estrada's schooling and the basics of where he had been employed (in the Justice Department)); Orrin Hatch, Abortion Stances Based in Religion, Monday, September 8, 2003, ROLL CALL, 2003 WL 7691833 (claiming that Pryor, his "highly qualified" candidate, was being blocked on religious grounds); Neil A. Lewis, GOP Senators Try to Change Filibuster Rules, Saturday, May 10, 2003, SAN DIEGO UNION-TRIBUNE, at A6, 2003 WL 6582851 (quoting Senate Republican leader Bill Frist as referring to both Miguel Estrada and Priscilla Owen as "highly qualified and intellectually superior" without explaining the basis for this characterization).

[4] *See* materials cited in note 3, *supra*.

[5] Bill Marshall and Michael Gerhardt tell us that we have overstated our point about the lack of concern about merit on the part of politicians. 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. 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.*

This Article presents a set of simple and objective measures to evaluate judicial merit, placing judges in a tournament of sorts. Our simple measures do not provide a perfect measure of judging skill. But that is not the standard at which we are shooting. The goal is to demonstrate the availability of a set of objective measures, on which data is easy to collect and analyze, that would do better than the current system to identify at the outset a merit-worthy pool of Supreme Court candidates. We are seeking to open the floor to a merit debate because 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. And where politics motivates the selection of a particular nominee, politicians will have less room to hide such politics behind merit-based justifications.

Some of our colleagues have raised the objection that merit and ideology are fundamentally inseparable in the context of judging. At some level, this may be right.[6] Nevertheless, there is value to attempting a separation at the extremes. The reason has to do with the claims that politicians make regarding candidates for judicial office. Politicians regularly claim that they are nominating (or confirming) based on merit and opposing the candidates of others based on ideology. Implicit therein is a claim that not only are merit and ideology separable, but that merit is good and ideology is bad. Our goal with the Tournament is to force the politicians to come clean on their motivations. Of course, they do not tell us what they think constitutes merit. But motivations are what we are attempting to flush out with the tournament. Hence, we came up with a set of simple measures that we think they would have difficulty denying are indicative of merit. Confronting the politicians with these, we hope, will force them to explain how they can be claiming someone as the most qualified if she does not do well on our simple measures. Maybe the end result will be that the politicians will be forced to acknowledge that there is not a distinction between what they mean by merit and ideology.[7] From our perspective, even that would be a positive result.

---

[6] Although, perhaps not entirely right. It has hard to see how a measure such as productivity, relating to how many opinions a judge writes, is fundamentally the same thing as a judge's ideologically driven views on abortion or affirmative action.

[7] 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 (2001) (drawing a distinction between an ideology of process and one of substance and arguing that it is the former that is legitimate).

The rhetoric about merit aside, and assuming that merit and ideology are separable, there is a long history of arguments that politics and ideology should matter in the nomination process.[8] 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 42 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. And, to the extent that the public concludes that one or the other side is being unduly ideological, it can penalize them in the next elections.

Some will see the search for a set of objective measures as a pointless exercise because they think that there is no way to measure or quantify what it means to be a good, let alone great, judge. We agree 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 construct measures that can be used to make a meaningful *relative* evaluation. 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 Cardozo's greatness as a judge to numbers. But one can look at how many times Armstrong has won the *Tour de*

---

[8] *See* Erwin Chemerinsky, Ideology and the Selection of Federal Judges, 36 U.C. DAVIS L. REV. 619 (2003). For more on the arguments that ideology should (and has) played a important role in the process, see 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 ideology was of primary concern with President Washington's nominee, John Rutledge. *See also*, James J. Brudney, Recalibrating Federal Judicial Independence, 64 OHIO ST. L.J. 149 (2003); John C. Eastman, The Limited Nature of the Senate's Advice and Consent Role, 36 U.C. DAVIS L. REV. 633 (2003); Ed R. Hayden, Judicial Selection: A Pragmatic Approach, 24 HARV. J.L. & PUB. POL'Y 467 (2003); Matthew D. Marcotte, 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); William G. Ross, The Role of Judicial Issues in Presidential Campaigns, 42 SANTA CLARA L. REV. 391 (2002); John S. Baker, Ideology and the Confirmation of Federal Judges, 43 S. TEX. L. REV. 177 (2001). 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, 132 (2001) (asserting that that nominees should not be asked about legal questions, because the purpose of the confirmation should not be to have a politically leaning court, but a fair court).

France and compare his numbers to those of his peers. Similarly, one can look at Cardozo's opinions and see how much they were cited by other judges, how often they were discussed in the law reviews, and how often they made their way into the casebooks. Cardozo's numbers on each of these measures can then be compared to those of his peers.[9] 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 does help us get 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 at least launches a discussion on how to develop a widely accepted set of objective criteria. Objective criteria may never have the ability to select the very "best" candidate for promotion to the Supreme Court. But if the process narrows the field of potential candidates to a substantial extent (say to the top ten judges selected from a variety of different objective criteria relative to their peers), 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, justify why the candidate in fact is the best despite failing to do well on the objective criteria (e.g., 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 (deciding cases and writing opinions explaining the

[9] Richard Posner, in his work on Justice Cardozo and Judge Learned Hand, looked at a set of such measures. *See* Cardozo: A Study in Reputation (1993); The Learned Hand Biography and the Question of Judicial Greatness, 104 YALE L. J. 511 (1994).

Some of our critics assert that Cardozo makes a bad example of relative excellence because the view held by many of his greatness derives from his performance at the state court level and not on the Supreme Court (where he had only a six-year stint). In more stark terms, the argument is that that the fact that 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 Cardozo makes it onto seven of them. *See* Lee Epstein et al. Rating the Justices: Lessons From Another Court (unpublished draft, presented at the Midwest Political Science Association Meeting, April 1992, on file with authors).

decisions) is often much the same as 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.[10] 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.[11] This Article takes on the harder question of how one might implement such a tournament.

Others have studied citations counts and more generally the question of which circuit court judge is the best.[12] 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 (in our case 1998 to 2000 for a total of 3 years) 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, does well in ours as well. However, he is no longer universally the best. For one of

[10] 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, Supreme Court decision making involves more policy making than the lower courts. In a similar vein, because there are nine justices and because of the tremendous importance of anything that the Court says, Supreme Court decision making involves a much greater need for negotiation and compromise than on a circuit court where the issues are often mundane. As well, a significant portion of a Supreme Court justice's job involves deciding the cases on which 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, policy making abilities, or certiorari denying skills. We do not see how it could.

[11] Stephen J. Choi & Mitu Gulati, A Tournament of Judges, 92 CAL. L. Rev. __ (forthcoming, 2004).

[12] *See* 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) (providing a study of the influence of circuit court judges sitting in 1992 with at least six years of tenure at the time based on citation counts for the opinions written by each judge from the beginning of the judge's tenure to 1995). *See also* Richard Posner, Is the Ninth Circuit Too Large? A Statistical Study of Judicial Quality, 29 *Journal of Legal Studies* 711 (2000) (providing an empirical test of whether the quality of the Ninth Circuit opinions is harmed by the large number of judges on the circuit based on Supreme Court reversals and citation counts).

our measures of opinion quality, the number of outside circuit citations garnered by a judge's top 20 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 and how often judges are cited, we are concerned with how independent a thinker is a particular judge.

Part II compares the current system to our tournament and uses that comparison as a base 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 faired 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 less the candidate's period of influence on the court), and candidates who were appointed to the bench by a President of the same political party as the current President. For the sake of simplicity, we use these elemental criteria to constitute the pool of available candidates. Plus, our quarrel is not with this first step. It is with the current process' next step.

In step two of the current process, as best we can tell, the President winnows 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.[13] 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 is

[13] By "current" tournament we mean the selection process being used by the current administration.

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 would be a reasonable debate on a candidate's ideology, such as whether she is a strict constructionist, is reduced to quibbling over her expected position on issues like affirmative action and abortion. That, in turn, means there is room for strategic judging. A lower court judge with aspirations to the High Court has an incentive to vote on the political issues in ways that will most please her potential sponsors (and, if possible, least offend her potential opponents). Thus, a judge with higher aspirations need not be productive—moving her docket along to promptly and efficiently resolve disputes—nor need she draft well-written opinions oft cited by his 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. The current selection process and its focus on highly political proxies of a candidate's ideology compromise the independence and productiveness 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' performance on three relatively apolitical factors: (1) productivity; (2) quality; and (3) independence. People will disagree with our methods of measuring performance because they are imperfect. We use a set of objective measures such as numbers of opinions written (as a proxy for productivity) and numbers of citations (as a proxy for quality). Yes, citations are a highly imperfect measure of quality in the same way that the SAT is a highly imperfect measure of aptitude for college. And like the SAT, our measure may even have embedded biases.

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

First, the measures we propose are relatively apolitical and objective. When a judge works harder, produces quality opinions, is cited by his colleagues, and makes decisions independently of his

[14] *See* the materials cited in note 3, *supra*.

political party, he 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 currently made 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 bad type of 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 to do 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. The incentives for a judge are to signal her political ideology by voting on certain hot button issues in a way that pleases her political masters. On its own, that is easy to do. For example, one simply votes a particular way on an abortion or death penalty case. The complication, however, is that the judge seeking advancement will have to be careful so as not to signal her ideology to the other side because of the danger that they will muster their resources to block her. The game playing that will go on, therefore, will be to provide stealth signals to one's political masters, while sending ambiguous or misleading signals to one's opponents.[15] The costs here will be in reduced transparency and added uncertainty. Candidates will be pushed up with relatively blank records, with the hope that they will turn out to support the ideologies of their sponsors. The result will be a high level of uncertainty in terms of 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.

In discussions comparing the Tournament to the current system, we are inevitably asked about the ABA's evaluation process: Why does one need a Tournament if we already have the ABA's own assessment of

[15] A concrete example of this might be a Republican judge avoiding 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.

merit? The easy answer is that the current administration has decided to not 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 non partisan 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 person's articles, other writings, speeches, and legal briefs.[18] From what the ABA reports, the inquiry is exhaustive and thorough. And in the past, both Presidents and Senators have attached significant weight to the rankings that the ABA reports (it ranks the candidates as either well qualified, qualified, or not qualified).[19]

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 will 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 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.[20] 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 the reputational bond works to guarantee the unbiased quality of the evaluation, that is fine. But, even if it worked at one point in time, it is not working anymore. The current administration sees the ABA as having a liberal bias. And

---

[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 OF RTS. J. 37, 37-44 (2001).

[17] *See* The ABA Standing Committee on Federal Judiciary: What it is and How it Works (1999), available at http://www/abanet.org/poladv/scfedjud.pdf

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

[19] *Id.* For more detail on the role that the ABA's evaluations have played historically, see, e.g., Henry J. Abraham, Justices, Presidents and Senators, 23-28 (revised ed. 1999); George Waters & John A. Stokey, Shaping America: The Politics of Supreme Court Appointments, 83-85, 108-112 (1995).

[20] *See* ABA report, *supra* note X, at 1-12 (reporting on the importance of "confidentiality" in the process and not providing the reader any idea of the precise evaluative tools that are going to be used).

there is some empirical evidence to back their suspicions.[21] The value of our Tournament in comparison to the ABA ranking system is that where the ABA evaluation process and eventual ranking is non-transparent and unverifiable, ours is designed to be precisely be those things. It may be that the ABA's rankings are better measures of merit (and we are not, by any means, advocate getting rid of them). But the ABA's criteria are of little use in separating the politicians' claim 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, but we think that there is far less of a danger than with an ABA type system.

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.[22] But the Constitution says nothing about what standards are to be used. The Appointments Clause provides that the president "shall nominate, and by 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 justice, but to prevent nepotism and tyranny.[24] In so

[21] *See* James Lingren, 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*, 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 (2001) (finding Lingren's methodology flawed and conclusions "remarkably overheated"); 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).

[22] *See* Baker, *supra* note X.

[23] U.S. CONST. art. II §2.

[24] Michael J. Gerhardt, Toward a Comprehensive Understanding of the Federal Appointments Process, 21 HARV. J.L. & PUB. POL'Y 467, 474-75 (1998) writes:

> Contrary 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

drafting, the Appointments Clause provides little guidance on candidates' ideal qualifications.[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 norm that it is acceptable to select (or block) a candidate on ideological grounds.[27]

> tyranny, whereas many others were concerned mostly with preventing monarchic despotism.

[25] *See* John S. Baker, Ideology and the Confirmation of Federal Judges, 43 S. TEX. L. REV. 177, 178 (2001), explains:

> The 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.

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 (2002) (looking at other countries' constitutionally-mandated criteria for judicial appointees and comparing results with the United States).

[26] Michael Gerhardt, Norm Theory and the Future of the Federal Appointments Process, 50 DUKE L. J. 1687 (2001). Norms include, for example, the practice of senatorial courtesy. 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. This practice has expanded from a state appointment to a circuit appointment, even though the origins of the practice are unclear. *See* Brannon P. Denning, The Judicial Confirmation Process and the Blue Slip, 85 JUDICATURE 218 (2002) (describing the blue slip norm).

[27] See James J. Brudney, Recalibrating Federal Judicial Independence, 64 OHIO ST. L.J. 149, 157 (2003) ("Yet while merit-based considerations are both necessary and important, they have not been viewed as sufficient by either of the two branches constitutionally charged with designating members of the federal bench."); David M. Levitan, The Effect of the Appointment of a Supreme Court Justice, 28 U. TOL. L. REV. 37, 69 (1997) (analyzing where the appointment of a single Supreme Court Justice has directly affected the law through reversals of earlier decisions and concluding that: "[I]t 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* John C. Eastman, The Limited Nature of the Senate's Advice and Consent Role, 36 U.C. DAVIS L. REV. 633, 647 (2003) (arguing that the Senate's confirmation exists only to prevent the President from selecting a nominee who does not possess due qualifications for office; essentially then, 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); *But see* Hearing before the Senate Committee on the Judiciary Subcommittee on Administrative Oversight and the Courts on should Ideology Matter?: Judicial Nominations 2001, June 26, 2001, Statement by Professor Cass Sunstein, 50 DRAKE L. REV. 463 (2002) (noting that today's judiciary is too active and needs to be checked).

In sum, there appears little room for an argument that the Constitution requires that the President and the Senate use standards in selecting judicial candidates. Indeed, given the open ended grant of discretion of the President (in nominating) and the Senate (in its advice and consent) it is likely that the imposition of standards would require a constitutional amendment—unless, of course, the President or the Senate were to voluntarily adopt standards.

We can imagine two "strong" forms and one "weak" form of our tournament proposal. Under the first strong form, the President chooses a nominee from only the pool of tournament winners. Under the second, the President has discretion to choose a nominee outside of the tournament winners. However, the President (and the members of the Senate) would not be permitted to introduce merit-based arguments outside of the tournament's factors. Instead, only non-merit-based, more ideological arguments would be allowable to justify a candidate outside of the tournament winners (forcing those who desire a particular candidate based on ideology to make this explicit). How would these strong forms of our proposal happen short of a constitutional amendment? We suspect that they wouldn't. The President and the Senate are unlikely to ever agree to relinquish their discretionary, constitutionally-derived power.[28]

The weak form of our tournament, however, does not present such insurmountable problems. This form calls for broad stakeholder participation–the media, civil society, and the broader public–and, further, the process itself is simple: information dissemination. 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 in fact is 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

[28] One scenario where they might agree to voluntarily restrict their power is where the level of conflict in the appointments process heightens even further and 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 Law, Appointing Federal Judges: The President, the Senate, and the Prisoner's Dilemma (draft dated 10/29/03, on file with authors) (modeling the dynamic between the President and the Senate in game theoretic terms).

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 the public's pressure, the President and Senate in turn, may view those who do better in the tournament with more favor.

In order for the weak form of the tournament to work, these stakeholders must exist. Further, the tournament must be sustainable and not default to hot button issues. Obviously, 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 widely available. The tournament introduces a higher degree of accountability, meaning that if an undeserving nominee is selected, the President and Senate will have to justify their decision. Thus, although the public may have difficulty understanding constitutional interpretation, they do understand that in an ideal world hard work is basis for a promotion. Thus, the tournament will attract new stakeholders.

The second necessary element for a merit based tournament is that it be sustainable, meaning that the process be constructed to hold up over time. One of the reasons that hot-buttons have become so important may be the public's rational 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've said before, shrouded in secrecy, and therefore 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, for example, gun rights and affirmative action. In addition, these also tend to be the issues that interest groups are most interested in. The result then 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 has fallen. The tournament will be sustainable because hot buttons are simply proxies for information that was formerly very costly.

III. Constructing the Tournament

We limit our sample to federal circuit court judges other than those on the federal circuit.[29] While, in theory (and, in the not too distant past, in practice) the President may select a Supreme Court nominee from the pool of state court judges and even from among non-judges, 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 parties—the assumption, however, being that the sitting President will focus on the top performers from his 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

[29] The two circuits that have the most specialized and, therefore, non-comparable 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 the D.C. Circuit had at least a meaningful portion of its docket that 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 Seventh or First Circuits. That said, a comparison of their respective numbers of citations to their top twenty opinions may be more meaningful. Focusing on only a judge's top 20 opinions reduces (if not eliminates) the advantage of judges writing 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 Edwards ranks 30th overall, but that he has high productivity and independence scores and low citation numbers, a discussion could occur as to whether citation scores are an inapplicable measure for him because of the D.C. Circuit's unusual docket. It might, for example, be argued that Edwards' relative ranking on our invocation scale or 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 (and transparent with respect to ideology) discussion than one initiated in a vacuum. On the historical evolution of the D.C. Circuit's Docket, see Christopher P. Banks, Judicial Politics in the D.C. Circuit Court (1999).

retiring or taking senior status suggests either a diminished capacity for, or interest in, judging. As well, we include only those judges who were appointed prior to January 1, 1998. In effect then, the tournament has a six year apprenticeship period (given that we calculate data up to May 31, 2003) before one can be considered for promotion. The apprenticeship period served the function of enabling 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 in the tournament 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 flat 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 quite possible (although 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 98 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
1	4
2	5
3	7
4	8
5	13
6	8
7	10
8	5
9	12
10	8
11	10
DC	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 objective measures that get at the quality and quantity of decision-making and opinion writing. We use measures that roughly fall 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 any given circuit judge hears is largely a function of the circuit she sits on. There are significant differences in the caseloads across the circuits, but the one commonality is that the burdens are overwhelming.[30] So much so that almost no judge can hope to provide a publication worthy statement of reasons in every case. Some

[30] 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, 687 (2001).

judges, however, provide published statements of reasons in more cases than others.

Presumably, it takes greater effort and skill to be able to write more published opinions.[31] Measuring effort exerted by a judge in the past 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 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.[32] Unpublished opinions represent opinions which 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 that we want to measure the willingness to exert high 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. While such pursuits may be valuable, we focus solely on published opinions for two 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.[33] 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

[31] *See* Alex Kozinski & Stephen Reinhardt, Please Don't Cite This!, CAL. LAW., June 2000 (explaining the greater effort that goes into writing published opinions); *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).

[32] The practice of issuing unpublished opinions in the Courts of Appeals began in the 1970s and today accounts for more than 75% of all decisions. Stephen L. Wasby, Unpublished Decisions in the Federal Courts of Appeals: Making the Decision to Publish, 3 J. App. Prac. & Process 325 (2001) (describing the Ninth Circuit's decision process to publish an opinion or not).

[33] *See* Fred R. Shapiro, The Most- Cited Legal Scholars, 29 J. LEGAL STUD. 409 (2000) (reporting Posner to be among the most cited of the jurisprudential giants, including those such as Oliver Wendell Holmes and Roscoe Pound). For research reporting Posner's opinion writing rate, see William M. Landes et al., Judicial Influence: A Citation Analysis of Federal Courts of Appeals Judges, 27 J. LEGAL STUD. 271 (1998).

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 in Table 2. The Appendix provides the ranking for the entire set of tournament judges (reported in Appendix Table B).

Table 2

Published Opinions Written from 1998-2000

(for the ten judges with the highest number of published opinions)

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

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 defined as those who are in the top 50% of judges in the entire sample (n=98) based on the number of published opinions (majority, concurrences, and dissents) (A). Bottom judges defined as those who are in the bottom 50% 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). Posner and Easterbrook, the two publishers of the largest numbers of opinions, are each more than three standard deviations above the mean. The next two highest, Flaum and Wood, are both greater than two standard deviations more than the mean. The mean number of published majority opinions is 83.6. As with the number of total published opinions, both Posner and 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 court cultures (the ethic on some courts may be to publish more opinions) and the court's caseload. [34] Circuits may exhibit different norms on when to publish an opinion. And, some circuits may have a norm of deciding more cases with less time (or they may face a greater or more complex case load). Judges from a low productivity circuit may well 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 distributions of circuits for top judges and the bottom judges (distinguished based on number of published opinions) are identical (χ^2

[34] A number of scholars have observed the importance of circuit norms in determining at least 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, 84-85 (2001); Ahmed Taha, Publish or Paris: Evidence on How Judges Allocate Their Time (forthcoming, AMER. L. & ECON. REV. (Fall 2003)). The Seventh Circuit, for example, appears to have a culture of publishing a high fraction of its opinions and the Third Circuit, in contrast, appear to have the opposite culture. *See* Mitu Gulati & Catherine M.A. McCauliff, On Not Making Law, 61 LAW & CONTEMP. PROBS. 157, 210-11 (1998). Along these lines, some readers will have noticed that the top scorers on both the number of published opinions and the number of published pages are from the Seventh Circuit (Wood, Posner, Easterbrook). But one should not be too quick to say that their high scores should be discounted because those scores are a function of their circuit's norms. This is because the norm, in turn, is a function of the inclinations of the individuals on the court. Indeed, it is quite likely that the emergence of the Seventh Circuit's norm is in large part due to the influence of Posner (and then Easterbrook and perhaps now 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 (2002). From this vantage point, it can be argued that the Posner, Easterbrook, and Wood should receive more credit, and not less, for being on a circuit that has a norm of high production.

= 34.697 (11 d.f.) (prob. ≤ 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.[35]

To control for the possible influence of circuit based norms on publication rates, we determine the mean number of opinions published for judges of each circuit. The mean number of total opinions for judges of the Seventh Circuit is equal to 185.2 (the highest among all the circuit means). The mean number of total opinions for the 3rd circuit, on the other hand, is equal to 60.1. For each circuit other than the Seventh Circuit, we calculate the number of opinions by which the other circuit's mean was less than the Seventh Circuit. 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. Samuel Alito, a judge on the Third Circuit, wrote 70 opinions from 1998 to 2000. We adjusted Alito's total opinion count upward to 195.1. After the adjustment, all the circuits have the same mean number of total number of opinions written. Any differences among judges will therefore be determined solely by each judge's standing relative to other judges within her own circuit. For example, if Morris Arnold is relatively far above his peers in the 8th circuit (compared with how much higher the most productive judges are above their peers in other circuits), 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).

[35] The ability to effectively utilize 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.

Table 3

Published Opinions Written from 1998-2000
Adjusted for Inter-Circuit Differences
(for the ten judges with the highest number of published opinions)

Judge	(A) Total Number of Published Opinions (adjusted for circuit variation)	Z-Score of (A)	(B) Circuit
Richard Posner	269	3.60**	7
Stephen Reinhardt	237	2.23**	9
Diarmuid O'Scannlain	234	2.10**	9
Frank Easterbrook	233	2.05**	7
Karen Nelson Moore	231	1.94	6
Ronald Lee Gilman	225	1.69	6
Gerald Bard Tjoflat	224	1.66	11
Jerry Smith	223	1.63	5
Paul V. Niemeyer	221	1.54	4
Dolores K. Sloviter	217	1.37	3

** 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 7th 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 defined as those who are in the top 50% of judges in the entire sample (n=98) based on the number of published opinions (majority, concurrences, and dissents) adjusted for intercircuit differences (A). Bottom judges defined as those who are in the bottom 50% of judges based on the number of published opinions (majority, concurrences, and dissents) adjusted for intercircuit differences (A).

The adjusted rankings capture solely variations within a circuit.[36] Even with the adjustment for intercircuit differences in means, note that

[36] The chi-squared test cannot reject the null hypothesis that the distribution of circuits in the top (representing the top 50% of judges based on the number of total published opinions correcting for intercircuit differences) and bottom groups of judges (the bottom 50%) is identical ($\chi^2 = 5.253$ (11 d.f.) ($p \leq 0.918$)).

Posner again is the highest scoring judge in terms of productivity (and the difference in Posner's production from the mean judge is statistically significant). After correcting for the greater mean number of opinions published in the Seventh Circuit, Posner's relative standing among other Seventh Circuit judges is high enough to place him ahead of high-scoring judges in other circuits. Easterbrook, however, drops to fourth place behind Reinhardt and O'Scannlain of the Ninth Circuit (all of whom are also significantly above the mean judge). Put another way, Reinhardt's relative ranking among Ninth Circuit judges is higher than Easterbrook's relative standing 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 us on exactly what those other justifications are.

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'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. What we have then 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 those opinions more. Examining customer use then provides a market test of the quality of a judge's opinions.[37]

[37] For a discussion of citation measures to rank judicial influence and the various problems using such measures, see Landes et al., *supra* note X at 271-76. Additional discussions of the limitations of using citation counts to measure judicial influence are in, Virgil Blake, Citation Studies--the Missing Background, 12 CARDOZO L. REV. 1961 (1991); Keith Ann Stiverson & Lynn Wishart, Citation Studies--Measuring Rods of Judicial Reputation? 12 CARDOZO L. REV. 1969 (1991). On citation analysis more generally, see Richard Posner, An Economic Analysis of the Use of Citations in

For each judge, we collected data from Sheppards (Lexis) on the number of times their 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 greater number of citations than other judges. Judges who are better liked, are on more respected circuits, who have been on the bench longer, have the favor of the current Supreme Court, etc., may receive more citations.[38] Judge Easterbrook, for example, has a history of being well-cited in many law periodicals and judicial opinions. The large number of citations Easterbrook has received in the past is due at least in part to the value of his analysis and the clarity of his presentation. Easterbrook's past then, makes it likely that his opinions during the 1998 to 2000 time period will be more likely to 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 pre-tournament inherent differences. Indeed, 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 (the Supreme Court, other circuit courts, district courts in other circuits, and state courts), and self-citations (where a judge cites herself). Some measures are more indicative of opinion quality than others. For example, citations by courts in other jurisdictions are more indicative of opinion quality than are citations from one's own jurisdiction. This is because the latter courts often have to cite opinions from the same circuit as binding precedent. For courts outside one's jurisdiction (termed

---

the Law 2 AMER. L. & ECON. REV. 381 (2000); *see also* Interpreting Legal Citations, Symposium Issue, 29 J. LEGAL STUD. (2000); Trends in Legal Citations and Scholarship, Symposium Issue, 71 CHI.-KENT L. REV. (1996).

[38] 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 does such an analysis, see Landes et al., *supra* note X.

"outside circuit citations"), however, citations occur only if citing the opinion serves a purpose in making an argument.[39]

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 is gathered from Sheppards on Lexis measured as of May 31, 2003. Posner and Easterbrook dominate (and are significantly different from the mean judge). The mean judge in the sample received 417.3 outside circuit citations. An indication of the domination is that both Posner and Easterbrook each have over four standard deviations more citations than the sample mean.

[39] *See* Landes et al., *supra* note X, 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.").

Table 4

Citations to Opinions Published from 1998-2000

(for the ten judges with the highest total number of outside citations)

Judge	(A) Total Outside Circuit Citations	Z-Score of normalized (A)	(B) SCT Citations	Z-Score of normalized (B)	(C) Law Review and Periodical Citations	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 L. Lynch	1023	1.99**	5	0.62	684	1.52	178	1.60	1
Bruce M. Selya	949	1.85	3	-0.04	727	1.65	364	2.28**	1
Paul J. 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
Diane Wood	678	1.20	3	-0.04	513	0.90	127	1.27	7
Harvie Wilkinson III	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 measured to 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 defined as those who are in the top 50% of judges in the entire sample (n=98) based on the number of outside circuit citations (A). Bottom judges defined as those who are in the bottom 50% 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 rated highest based on the number of outside circuit citations (collected from Sheppards on Lexis). Note that Posner and Easterbrook again receive the highest two rankings for both U.S. Supreme Court citations and law review and periodical citations.

Landes, Lessig, and Solimine suggest that judges who write more of their opinions typically engage in more self-citations (due to their greater familiarity with their own self-authored opinions).[40] Judges who add a large amount of value in their opinions consistently over several years typically will minimize the input of their clerks. In other words, judges who rely heavily on their clerks will produce opinions of more varying quality over the years due to the influence 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, therefore, that judges who have a larger number of self-citations (indicating more self-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 Wilkinson). Posner and 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 to decide a case with a minimal explanation (and sometimes not even that).[41] Other things equal, a judge with 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 distributions 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.) (prob. ≤ 0.001)).[42]

[40] *See* Landes et al, *supra* note X, at 274.
[41] *See* Gulati & McCauliff, *supra* note X.
[42] Consistent with our finding of circuit specific differences in outside circuit citations, Landes et al. also report that the Seventh and First Circuits do particularly well in their citation study. *See* Landes et al., *supra* note X, at 301-302.

Table 5

Outside Circuit Citations to Opinions Controlling for Total Number of Opinions Published from 1998-2000

(ten judges with the highest number of outside circuit citations to each judge's top twenty opinions)

Judge	(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 L. Lynch	734	2.56**	9.03	1.73	1
Frank Easterbrook	667	2.33**	6.25	0.73	7
Paul J. Kelly	654	2.28**	9.85	1.97**	10
Richard Posner	570	1.95	5.49	0.37	7
Bruce M. Selya	516	1.71	6.50	0.83	1
Anthony J. Scirica	496	1.61	14.50	3.04**	3
Frank M. Hull	460	1.43	10.90	2.25**	11
Karen J. Williams	455	1.40	11.02	2.28**	4
Edward Earl Carnes	444	1.34	8.92	1.70	11
Harvie Wilkinson III	425	1.24	7.64	1.28	4

* Indicates a Z-Score of 1.96 or higher (representing a two-sided probability of <5% for a normal distribution). Outside circuit citations measured to 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 20 citation-receiving cases (A): $\chi^2 = 15.466$ (11 d.f.) ($p \leq 0.169$). Top judges defined as those who are in the top 50% of judges in the entire sample (n=98) based on the number of outside circuit citations to the top 20 citation-receiving cases (A). Bottom judges defined as those who are in the bottom 50% of judges based on the number of outside circuit citations to the top 20 citation-receiving cases (A).

One possible correction is to look at the average citations per majority opinion. These numbers are reported on column (B) of Table 5. The problem with average numbers of citations per majority opinion is that they give less credit to a judge who writes several outstanding opinions and adds on many other smaller opinions compared to a judge who just writes the same number of outstanding 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 more smaller opinions, we look at the total outside circuit citations to each judges' top 20 citation receiving opinions. So regardless of the fact that Posner wrote far many more opinions than Sandra Lynch during the 1998-2000 time period, we sum only the outside circuit citations to each judge's top 20 citation receiving opinions. Column (A) of Table 5 reports the Top twenty opinion outside circuit citation count for the ten highest-scoring judges. (Appendix Table E reports on the entire sample judges).

Looking at the same number of opinions for each judge helps control 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 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 20 opinions) and the bottom judges ($\chi^2 = 15.466$ (11 d.f.) ($p \leq 0.169$)). In addition, given that judges on the Supreme Court write far fewer opinions than those on the lower courts (and it is a promotion to the High Court that we are looking at these numbers for), an argument exists that it is the judge's performance on her best opinions that we should look at and not just her average performance on the full set of opinions.

Here the two highest scoring judges for the top twenty opinion citation count are Lynch of the First Circuit and Easterbrook. Posner, while still in the top 5 is no longer the highest scoring judge. In terms of approach, one might infer that Lynch focuses more on crafting high citation opinions than Posner. Posner's individual opinions are not as heavily cited as Lynch's opinions. However, Posner's sheer productivity results in more overall citations than Lynch.

The use of citations will strike some as outrageous. One reason for this is that many believe that it is outrageousness, as opposed to quality, that gets 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 citations for authority as opposed to citations where the judge is attempting to distance herself from someone else's outrageous statements. Along those lines, a finer tuned measure could also eliminate negative citations, that is, citations of disagreement with a position.[43] Finally, to the extent it appears that law journals articles are more likely to cite outrageous opinions, those numbers can be discounted (and are not included in the citations measures in this paper aside from Column (C) in Table 4).

To test for 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 in the Lexis Sheppard's service) for the top 20 outside circuit citation-receiving opinions for each judge. In comparison, we obtained the number of negative citations to the top 20 citation-receiving opinions for the median judge as well as the immediate 5 judges above and below the median (for a total of 11 comparison judges). Table 6 reports the negative citation counts.

[43] One of our colleagues who is against using citation measures says that no one will ever construct the finer tuned measure because they are too difficult to do. We disagree. If people find the idea of using such data attractive (and we concede that that is a big if), that 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 Count for Top 20 Citation-Receiving Opinions

Panel A: Negative Citation Count for Judges with Highest Top 20 Citation Counts

Judges with Highest Top 20 Outside Circuit Citation Count	Outside Circuit Citations to Top 20 Opinions	Negative Outside Cir. Citations to Top 20 Opinions	Fraction of Negative Outside Cir. Citations
Sandra L. Lynch	734	67	0.091
Frank Easterbrook	667	79	0.118
Paul J. Kelly	654	22	0.034
Richard Posner	570	65	0.114
Bruce M. Selya	516	33	0.064
Anthony J. Scirica	496	39	0.079
Frank M. Hull	460	72	0.157
Karen J. Williams	455	17	0.037
Edward Earl Carnes	444	38	0.086
J. Harvie Wilkinson III	425	48	0.113

Panel B: Negative Citation Count for 11 Judges Centered on the Judge with the Median Number of Outside Circuit Citations for the Top 20 Opinions

Median Judges	Outside Circuit Citations to Top 20 Opinions	Negative Outside Cir. Citations to Top 20 Opinions	Fraction of Negative Outside Cir. Citations
David S. Tatel	265	62	0.234
Edith Jones	262	20	0.076
James B. 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 R. Murphy	253	24	0.095
Eugene W. Davis	250	25	0.100
Alice M. Batchelder	250	23	0.092
Diana E. Murphy	245	15	0.061
Carlos F. Lucero	240	13	0.054

The fraction of negative citations for the judges receiving the highest number of outside circuit citations for their top 20 opinions is equal to 0.089. In comparison, the median judges had a 0.106 proportion of negative citations in the total number of outside circuit citations to their top 20 opinions. 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 proportions is not statistically significant (t-statistic = -0.833).

Another concern with the use of citation counts is that more senior judges (including the chief judge of a circuit in particular) have the ability to assign themselves the "choice" opinions from the panels on which they serve. Seniority, independent of the inherent ability of a particular judge, therefore may drive higher citation counts. To assess the importance of seniority and chief judge status, we again compared the top 10 judges based on the top 20 citation-receiving opinions against the 11 median judges. We also compared both against the bottom 10 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 10 judges (Posner and Wilkinson) equals the number of chief judges in the bottom 10 judges group (Edwards and Martin).

So which measure 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 Cir. Citations	**Total US SCT Citations**	**Law Review and Periodical Citations**	**Outside Cir. Citations to Top 20 Opinions**
Total Outside Cir. Citations	1			
Total US SCT Citations	0.589	1		
Law Review and Periodical Citations	0.827	0.571	1	
Outside Cir. Citations to Top 20 Opinions	0.897	0.486	0.690	1

Aside from looking at citations, another method exists of measuring a judge's perceived quality. Standard practice in citing to a judge's opinion is to cite to the 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: "See, Judge Fancy said much the same in her opinion." The key element here is that the reputation of Judge Fancy is being invoked 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 judge's authority and influence among her peers than do citation rates alone.[44] 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 overall reputation.

To maintain a relatively level contest, we focus solely on invocations related to opinions authored by each judge in the tournament from 1998 to 2000. Easterbrook, for example, may well have a large number of invocations relating to his opinions written prior to 1998. We, however, set Easterbrook's past invocation count to zero and focus solely on how many times courts invoked 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 up to May 31, 2003 on Westlaw. Invocations include reference to a specific judge's name (related to an opinion in the sample time period) in both the text and a parenthetical with one exception. Opinions which 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.[45]

---

[44] 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, 224 (2001); David Klein & Darby Morrisroe, The Prestige and Influence of Individual Judges on the U.S. Courts of Appeals, 28 J. LEGAL STUD. 371, 372 (1999).

[45] 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 is a signal that the cited opinion is viewed as special.

Table 8

Invocations to Opinions Published from 1998-2000

(ten judges with greatest total invocations)

Judge	(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%	7
Frank Easterbrook	103	3.36**	0.44	4.67**	99.0%	7
Guido Calabresi	23	1.85	0.23	2.35**	91.3%	2
Harvie Wilkinson III	19	1.66	0.18	1.73	73.7%	4
Michael Boudin	13	1.30	0.10	0.70	84.6%	1
Patrick E. Higginbotham	12	1.22	0.12	0.96	41.7%	5
Diarmuid O'Scannlain	11	1.14	0.08	0.43	54.5%	9
Edith Jones	11	1.14	0.11	0.83	81.8%	5
Diane Wood	10	1.05	0.05	0.01	20.0%	7
Michael J. Luttig	10	1.05	0.12	0.96	80.0%	4

** 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 defined as those who are in the top 50% of judges in the entire sample (n=98) based on the number of invocations (A). Bottom judges defined as those who are in the bottom 50% 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 F reports invocation rates for the entire sample of judges).

The Posner-Easterbrook dominance on invocation numbers is dramatic. They each have at least 5 times the invocations of the next highest judge. The mean number of invocations for the sample is 6.8 with a standard deviation of 20.4. Both Posner and Easterbrook have at least 5 times the standard deviation number of invocations compared with the mean judge. Interestingly, 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) (χ^2 = 13.863 (11 d.f.) (prob. $\leq$ 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 Posner and Easterbrook's invocations that are to their majority opinions. For them, it is close to 100%. For almost everyone else (except Calabresi), a significant fraction of their invocations are to their dissents and concurrences. At the low end, for example, only 20% of 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 invest their skill and attention to such opinions (leading to more invocations). Posner, Easterbrook, and Calabresi appear, in contrast, to 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 we most care about in our judges. It is largely for reasons of independence that we give federal judges lifetime appointments and virtually iron clad job security. It is also for this reason that many of us feel disappointment when we see Supreme Court votes fall according to the known political sympathies of the judges.[46]

[46] The votes in Bush v. Gore being a prime example of the type of voting pattern that caused great disappointment and produced heated accusations of partisanship. *See, e.g.,*

A different concern about judicial independence stems from the nature of multi-member 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 and celebrated (regardless of the political parties of the opposing judges). And, more important, this is a trait that is 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 a concurrence forces the writer of the majority opinion to exert greater effort on his opinion than what he would have exerted otherwise. The real extra effort however, is what the judge writing the independent opinion exerts. The reason for this is that the judge does not generally receive a break on the number of majority opinions she is assigned if she writes more dissents and concurrences. Writing those dissents and concurrences is additional work that the judge has to do, over and above her regular load of assigned majority opinions.[47] In sum, the number of dissents and concurrences one writes provides not only a measure of a judge's willingness to annoy her colleagues, but also a measure of her 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 in question

---

637 Law Professors Say By Stopping the Vote Count in Florida, The U.S. Supreme Court Used Its Power To Act as Political Partisans, Not Judges of a Court of Law, N.Y. TIMES A7 (Jan. 13, 2000) (advertisement); Jack M. Balkin, Bush v. Gore and the Boundary Between Law and Politics, 110 YALE L. J. 1407, 1408 (2001); Jeffrey Rosen, The Supreme Court Commits Suicide, NEW REPUBLIC 18 (Dec. 25, 2000), Mary McGrory, Supreme Travesty of Justice, WASH. POST A3 (Dec. 14, 2000; Linda Greenhouse, The Court's Credibility at Risk, N.Y. TIMES (Dec. 11, 2000).

[47] Our understanding is that this is an informal norm on the federal circuit courts of appeals. We do not, however, have formal support for the point.

is likely to have a strong ideology (one that she is unwilling to compromise on). A person with these views could simply place a negative weight on the number of independent opinions written and run the tournament that way. The important thing is not who wins the tournament, but that there be transparency regarding the promotion criteria being used.[48]

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.[49] Here, we use data on both the dissents that a judge writes and the dissents that are written against her. We classify each judge's opposing opinions as either against a judge of the same party or not (dissents against unsigned opinions are not counted). Where more than one judge is on the opposing side, we focus on the opinion writer (and if multiple opinions 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 the opposing opinions where 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 in terms of opposing a judge of the same party (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 only active for part of the time period

[48] One question we had 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 the Ninth and the Third. 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.

[49] There is a large 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). For more detail on the research, see Brudney, *supra* note X.

(e.g., just for 1998) we gave a correspondingly lower weight (e.g., weight of 1).

For example, Posner on the Seventh Circuit faced a mixture of judges consisting of 75% Republican-nominated judges. Even were Posner (a Republican-nominated judge) simply to oppose other judges regardless of political affiliation, we would expect that he would oppose other Republican-nominated judges 75% of the time. Call the baseline percentage of Republican-nominated judges on the circuit the predicted same party opposing rate for Republican 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, he would score a -0.25 under our measure, dissenting against same party judges 25 percentage points less than an independent acting judge would. The more negative the independence score, the more aligned a particular judge is with their party line. On the other hand, if the Republican-nominated judge on the Seventh Circuit dissented against other Republicans 95% of the time, she would score a +0.20, being 20 percentage points more likely to dissent against a judge of the same political party. Table 9 reports the two measures of independence.[50] (Appendix Table G reports the independence measures for the entire sample of judges).

[50] 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. We are in the process of collecting that data for a separate article. *Cf.* Cass Sunstein, David Schkade, and Lisa Michelle Ellman, Ideological Voting on the Federal Courts of Appeals: A Preliminary Analysis, (unpublished draft available on SSRN.com (abstract_id #442480)) (reporting a higher likelihood of voting along party lines on certain key issues like abortion).

Table 9

Number of Opposing Opinions and Independence Ratings for the period 1998-2000

(twenty judges with best independence ratings)

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

** 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 9th 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 defined as those who are in the top 50% of judges in the entire

sample (n=98) based on the independence rating (A). Bottom judges defined as those who are in the bottom 50% 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 Posner is still in the top 10 (and Easterbrook in the top 20), other judges who scored well in terms of productivity and quality are not ranked as highly in terms of independence. Reinhardt, for example, wrote the second highest inter-circuit adjusted number of opinions. Reinhardt, however, is not among the top 20 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 opinions writers in the top 20 independent judges are in the bottom two-thirds of the list. Edith Jones, for example, wrote 21 dissents and concurrences (tied for second highest on the list) and is the last person on the list in terms of independence ranking.[51]

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 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. The reason is that closeness to zero suggests even handedness. That is, neither a preference nor an aversion to others appointed by a President of the same part. In contrast, a high positive score, while suggesting a change in ideology, still suggests an ideology 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.

[51] The chi-squared test in Table 9 rejects the null hypothesis that the distributions of circuits among the top and bottom judges (divided based on their independence rating) are identical ($\chi^2 = 23.110$ (11 d.f.) ($p \leq 0.017$)). While 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.

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

Table 10

Ranking of Independence for the period 1998-2000 among "Independent Judges" (-0.100 to +0.100 range of independence rating)

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

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's 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 65 years old (measured as of 2003). While including all active judges in our sample allows us to rank judges based on relative performance with peers, in practice, we suspect that the President will select a Supreme Court nominee younger than 65.[52]

A. Do We Need to Pick a Winner?

One answer for the question of who should win the tournament 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.

[52] The reason being that a younger appointee will likely have a longer tenure (and, therefore, greater influence).

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	3rd	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 Posner and 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 2 finishers including Posner, Easterbrook, Lynch, Ebel, Luttig, Barkett, and Arnold.

Not having one winner has an advantage. 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 may focus on publishing as many opinions as possible regardless of quality or of independence. Having the President select among a pool of winners reduces the benefit to judges from gaming 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 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 their 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, the task becomes more difficult for a President to make the "best" claim, flushing out more explicit political explanations for the selection of a particular judge.

B. Composite Measures

More is possible, nonetheless, than simply picking multiple numbers of 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

[53] In the weak form of the tournament, this pressure arises solely from public attention to the extent the objective factors of the tournament are easily transmitted and absorbed by the public. *See supra* Part II (discussing the weak form of the tournament).

measure of quality and quality is a key element of the tournament. 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. A constraint on such behavior though is that it uses up resources that could be used to write other opinions. And, to the extent total number of opinions written is used as 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 (e.g., 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 10 times more opinions than the next highest judge this should get more weight than if Judge X writes only 1.1 times the number of opinions. Second, the ordinal ranking scheme gives equal weights to the Productivity, Quality, and Independence rankings. However, as a matter of policy and preference it is unclear why these various measures should receive equal weighting.

We therefore eschew ordinal ranking schemes. Instead, we rank judges based on a composite of the Productivity, Quality, and Independence measures. For the Productivity, Quality, and Independence measures we use the following:

Productivity: Number of Opinions (majority, dissenting, and concurring) for 1998-2000 adjusted for InterCircuit differences

[54] Interestingly, Landes et. al report from 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 X, at 302.

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

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 these terms makes it unlikely that this will make a significant difference

Indep.: 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:

Score = xProductivity + yQuality – xIndependence

Such that

$x + y + z = 1$

Table 12 reports the results of select weightings for all judges in the sample excluding those 65 and over.

Before proceeding, however, a few caveats are worth noting.

(a) 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. What this means is 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 subject to superstar effects such as citations and invocations (that make up our quality measure). In part, our use of log transformations of the number of citations to the top 20 opinions and the number of invocations helps reduce the skewness in the distribution due to the superstar effect (resulting in a more normal distribution of judges).

(b) *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, Carl Stewart receives a high independence score (e.g., close to zero). He however had only two dissenting and concurring opinions in the three-year period. 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. More refined versions of the tournament may nonetheless wish to take into account opposing opinion productivity in addition to political independence.

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

Table 12

Linear Combinations of Quality, Productivity, and Independence Measures for Judges Under 65 with Various Weightings (Quality, Productivity, Independence)

[Quality]

Posner
Easterbrook
(1,0,0)

Posner
Easterbrook
(.75,.25,0)

Posner
Easterbrook
(.75,0,.25)

Posner
Easterbrook
(.67,.16,.16)

Posner
Easterbrook
(.5,.5,0)

Posner
Easterbrook
(.5,0..5)

Posner
Easterbrook
(.42,.42,.16)

Posner
Easterbrook
(.42,.16,.42)

Posner
Easterbrook
(.33,.33,.33)

Posner
Easterbrook
(.25,.75,0)

Posner
Easterbrook
(.25,0,.75)

Posner
Easterbrook
(.16,.67,.16)

Posner
Easterbrook
(.16,.16,.67)

Posner
Easterbrook
(.16,.42,.42)

Posner
Easterbrook
(0,1,0)
[Productivity]

Posner
Easterbrook
(0,.75,.25)

Posner
Easterbrook
(0,.5,.5)

Posner
Ebel
(0,.25,.75)

Ebel
Luttig
(0,0,1)
[Independence]

(c) Critics have questioned our measures on the grounds that we have no evidence that they predict high quality Supreme Court performance. The implicit suggestion is that we first do an analysis of relative performance for current (and past) justices on the Supreme Court and then work backwards 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 on the question is that such an analysis would be difficult, although probably not impossible. Among the problems that arise here is that there are only nine justices at any one time, creating a problem of adequate numbers. We can imagine a study, however, that adjusts for time differences (caseloads and such) and compares justices across the decades.[58]

There are dozens of other caveats that could be added. The nature of these measures is that they are imperfect proxies. 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 in Tables 12 is the dominance of Posner across a wide range of weights. In the rankings, Posner wins in every possible ranking except for where Independence is given 100% of the weight (Ebel wins instead). Easterbrook likewise wins nearly as many second-place finishes. People may care about a number of different criteria. Nonetheless, the conflict among criteria may be more illusionary than real. At least in the case of Posner, one judge does seem to dominate across a range of different measures regardless of the weighting.

Of course, in applying a tournament of judges over time, there may not be a judge like Posner. The almost near domination of Easterbrook as the second place judge, however, stands as a counterexample to this possibility. Nonetheless, to examine how the tournament would play out without either Posner or Easterbrook, we ranked judges according to various weightings on the Quality,

---

[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. Rating the Justices: Lessons From Another Court (unpublished draft, presented at the Midwest Political Science Association Meeting, April 1992, on file with authors). This study did not, however, use the ranking to come up with a set of criteria that could help predict performance.

[58] Steve Goldberg had the interesting suggestion that our data could be used to work backwards 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.

Productivity, and Independence measures without them as reported in Table 13.

Table 13

Linear Combinations of Quality, Productivity, and Independence Measures for Judges Under 65 with Various Weightings Excluding Posner and Easterbrook (Quality, Productivity, Independence)

[Quality]
S. Lynch
Wilkinson
(1,0,0)

Wilkinson
S. Lynch
(.75,.25,0)

S. Lynch
Wilkinson
(.75,0,.25)

Wilkinson
S. Lynch
(.67,.16,.16)

Wilkinson
K. Moore
(.5,.5,0)

S. Lynch
Luttig
(.5,0..5)

Wilkinson
Niemeyer
(.42,.42,.16)

S. Lynch
Wilkinson
(.42,.16,.42)

Wilkinson
Niemeyer
(.33,.33,.33)

K. Moore
Gilman
(.25,.75,0)

K. Moore
J. Smith
(.16,.67,.16)

Ebel
Luttig
(.16,.16,.67)

S. Lynch
Luttig
(.25,0,.75)

Ebel
J. Smith
(.16,.42,.42)

K. Moore
Gilman
(0,1,0)
[Productivity]

J. Smith
K. Moore
(0,.75,.25)

Ebel
Smith
(0,.5,.5)

Ebel
Carnes
(0,.25,.75)

Ebel
Luttig
(0,0,1)
[Independence]

Without Posner or Easterbrook, no one winner emerges. However, several judges do come to the forefront, including most notably Wilkinson and Lynch (with the two highest numbers of 1st and 2nd 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 tradeoffs among the criteria.

C. And What about the Low Scorers?

Pointing out the low scorers may help incentivize circuit court judges to produce more and higher quality opinions (while avoiding simply towing the party line). As well, identifying low scorers may help eliminate nominees that a President may otherwise be tempted to put forth to the Senate (or alternatively 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, where a couple of 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, where Posner and Easterbrook dominate across the different combinations of weights. Table 14 reports the 20 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

20 Lowest Scoring Judges (Under 65 Years) Using Equal Weighting of Quality, Productivity, and Independence

20 Lowest Scoring Judges (Ranked Lowest to Highest) based on equal weighting of Quality, Productivity, and Independence	(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	11
Pamela Ann Rymer	2	96[q]	9	9
Martha Daughtrey	3	95[p]	5	6
Michael Daly Hawkins	4	93[q]	4	9
Alice M. Batchelder	5	92[i]	7	6
Karen Henderson	6	98[q]	8	DC
Robert H. Henry	7	91[p]	4	10
Rhesa Barksdale	8	85[p]	8	5
M. Blane Michael	9	90[i]	5	4
Carlos F. Lucero	10	94[i]	3	10
Deanell Reece Tacha	11	93[i]	13	10
Joel Fredrick Dubina	12	91[q]	8	11
Susan Harrell Black	13	78[q]	6	11
James B. Loken	14	97[i]	8	8
R. Guy Cole, Jr.	15	83[i]	3	6
David Bryan Sentelle	16	94[q]	11	DC
Merrick B. Garland	17	95[q]	1	DC
Alex Kozinski	18	86[p]	13	9
Cabranes, Jose	19	89[i]	4	2
Manion, Daniel	20	98[p]	12	7

[q] indicates that Quality was the worst ranking criterion; [p] indicates that Productivity is the worst ranking criterion; [i] indicates that Independence is the worst ranking criterion.

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

D. The Effect of Experience

One criticism of our methodology is that it may give undue weight to those judges with more experience on the bench. In comparing the 20 lowest scoring judges against the judges who perform the best in the tournament, however, it is unclear whether there in fact is 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 place finish in the composite measures (excluding Easterbrook or Posner):

Table 15
Years of Experience for the Top Judges in the Tournament

Top Judges in the Tournament	Years as a Circuit Court Judge (at start of 1998)
Barkett	4
Carnes	6
Ebel	11
Gilman	1
J. Smith	7
K. Moore	3
Luttig	7
M. Arnold	6
Niemeyer	8
S. Lynch	3
Wilkinson	14

Top judges are defined to include the top 2 winners of each individual measure of merit as well as those receiving either a first or second place finish in the composite measures (excluding Posner or Easterbrook)

The mean experience for the bottom 20 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 top judges compared with the non-top judges is not statistically significant from zero (t-statistic = -0.205)**.** Even when Posner and 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-statistic = 0.667).

We suspect that a fuller study might reveal that there is something of an age-experience life cycle profile, where 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 20 citation-receiving decisions. As discussed above, this measure controls for differences in publication norms across circuits (by looking at the same number of opinions, 20, for each judge).[60] Similarly, productivity is based on the total number of published opinions 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 of judges determined through the equal weighted composite measure.

[59] For a study along these lines (that uses invocations as a measure of the quality of output generated), see Mita Bhattacharya & Russell Smyth, How Fast do Old Judges Slow Down? 23 INT'L REV. L. & ECON. 141 (2003).
[60] *See supra* Part III.B.
[61] *See supra* Part III.A.
[62] *See supra* note X.

Table 16

Circuit Breakdown of Highest and Lowest Scoring Judges (based on the equal weighted composite measure of merit)

Circuit	Num. of Top Judges	Percent of Circuit	Num. of Bottom Judges	Percent of Circuit
1	2	100.0%	0	0.0%
2	2	50.0%	2	50.0%
3	5	100.0%	0	0.0%
4	5	71.4%	2	28.6%
5	6	75.0%	2	25.0%
6	4	57.1%	3	42.9%
7	4	50.0%	4	50.0%
8	1	50.0%	1	50.0%
9	2	28.6%	5	71.4%
10	3	37.5%	5	62.5%
11	3	37.5%	5	62.5%
DC	0	0.0%	8	100.0%
Total	37		37	

Equal Weighted Composite Measure = 0.333Quality + 0.333Productivity + 0.333Independence. Top judges defined as those who score in the top 50% of judges based on the equal weighted composite measure of merit. Bottom judges defined as those who score in the bottom 50% 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 in the top 50% of judges. Interestingly, the Seventh Circuit (home to Posner and Easterbrook) only has half of their judges in the top half of judges. The DC Circuit, in contrast, performs surprisingly poorly, with all of their 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 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 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 is therefore at least partly a reflection that 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. As well, 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 attention if put on the Supreme Court.

While judges from certain circuits do poorly in our tournament, therefore, we are hesitant to introduce intercircuit controls beyond those we have already employed (looking at only adjusted numbers of published opinions and the citations to only the top 20 opinions for each judge). Further, although this is a question that 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: Harvie Wilkinson, Michael Luttig, Edith Jones, Samuel Alito, and Emilio Garza (the "Bush Five").[65] The

---

[63] See, e.g., Landes at al., *supra* note X, 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, Saturday, May 31, 2003, 2003 WL 55332005 THE HARTFORD COURANT at A1; Stuart Taylor, Jr., Bush and the Supreme Court: Place Your Bets, November 16, 2002, NAT'L L. J., 2002 WL 26794497.

In addition to the five names mentioned in the text, two other names that are often mentioned are Alberto Gonzalez, White House Counsel and former Texas Supreme Court Justice, and Janice Roger Brown, a California Supreme Court Justice. A fuller study could compare the relative performances of the state court justices and see how Gonzalez and Brown stand against their peers as well as federal 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 Janice Roger Brown has been nominated to the D.C. Circuit, supposedly, in the minds of many, as a first step towards moving her to the U.S. Supreme Court. *See* Jim Puzzanghera, U.S. Senate Showdown Set on Stalled Judicial Nominees, San Jose Mercury News, November 7, 2003, Section A; Carolyn Lochhead, Democrats Intend to Block Vote on

question 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 Posner and Easterbrook to the side, Wilkinson and Luttig emerge as among the top performers (along with Lynch, Moore, Ebel, and Smith). More surprising, was to see that three of the Bush Five (Luttig, Alito, and Jones) have among the highest scores on the independence measure.

VI. Conclusion

So what do the numbers from the tournament tell us? Should Posner and Easterbrook be the President's nominees? Not necessarily. There may well be finer tuned measures that do not have them as tournament winners. Plus, 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 produce transparency from those with the power to decide on nominees. Posner and Easterbrook are useful in this analysis in that their numbers give us a basis to challenge the President's slate with the question: Why not them? Ideally, the President and the Senate would be able to answer with a critique of our measures and an explanation of why under the appropriate measures it would be someone else who would emerge as the leader.

Where the tournament adds value is in creating a de facto presumption; one that the President has to rebut (or else face public pressure to the extent 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). But the President has to justify his choice to the public. If he introduces someone other than Posner or 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

---

Bush Pick/Feinstein Opposes Nomination of State Justice to Appeals Court, The San Francisco Chronicle, November 7, 2003.

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 she 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, Posner or Easterbrook was not his candidate. Maybe Posner is too unpredictable and won't 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 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 Posner and the burden of justification will move to the Democratic senators to show why 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, we set the goal for our tournament as providing 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 at times in the past. Our view is that they should; that they 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.

Appendix

Table A
Federal Circuit Court Judges in the Tournament

Judge	(A) Circuit	(B) Year Appointed	(C) Age in 2003
Juan R. Torruella	1	1984	70
Bruce M. Selya	1	1986	69
Boudin, Michael	1	1992	64
Sandra L. Lynch	1	1995	57
Walker, John M.	2	1989	63
Jacobs, Dennis	2	1992	65
Calabresi, Guido	2	1994	71
Cabranes, Jose	2	1994	63
Parker, Fred I.	2	1994	65
Sloviter, Dolores K.	3	1979	71
Scirica, Anthony J.	3	1987	63
Nygaard, Richard L.	3	1988	63
Alito, Samuel A.	3	1990	53
Roth, Jane R.	3	1991	68
McKee, Theodore A.	3	1994	56
Rendell, Marjorie	3	1997	56
H. Emory Widener, Jr.	4	1972	80
Wilkinson III, J. Harvie	4	1984	59
William W. Wilkins	4	1986	61
Paul V. Niemeyer	4	1990	62
Luttig, J. Michael	4	1991	49
Karen J. Williams	4	1992	52
M. Blane Michael	4	1993	60
Diana Gribbon Motz	4	1994	60
King, Carolyn Dineen	5	1979	65
Higginbotham, Patrick E.	5	1982	65
Jolly, E. Grady	5	1982	66
Davis, W. Eugene	5	1983	67
Jones, Edith	5	1985	54
Smith, Jerry	5	1987	57
Wiener, Jacques L. Jr.	5	1990	69
Barksdale, Rhesa H.	5	1990	59
DeMoss, Harold R. Jr.	5	1991	73
Garza, Emilio	5	1991	56
Stewart, Carl E.	5	1994	53
Benavides, Fortunato	5	1994	56
Dennis, James L.	5	1995	67

Table A Continued
Federal Circuit Court Judges in the Tournament

Judge	(A) Circuit	(B) Year Appointed	(C) Age in 2003
Boyce F. Martin, Jr.	6	1979	68
Boggs, Danny	6	1986	59
Alice M. Batchelder	6	1991	59
Martha Craig Daughtrey	6	1993	61
Karen Nelson Moore	6	1995	55
R. Guy Cole, Jr.	6	1995	52
Ronald Lee Gilman	6	1997	61
Eric L. Clay	6	1997	55
Posner, Richard	7	1981	64
Coffey, John	7	1982	81
Flaum, Joel	7	1983	67
Easterbrook, Frank	7	1985	55
Ripple, Kenneth	7	1985	60
Manion, Daniel	7	1986	61
Kanne, Michael	7	1987	65
Rovner, Ilana	7	1992	65
Wood, Diane	7	1995	53
Evans, Terrence	7	1995	63
Pasco M. Bowman	8	1983	70
Roger L. Wollman	8	1985	69
James B. Loken	8	1990	63
Morris S. Arnold	8	1992	62
Diana E. Murphy	8	1994	69
Pregerson, Harry	9	1979	80
Schroeder, Mary M	9	1979	63
Reinhardt, Stephen	9	1980	72
Kozinski, Alex	9	1985	53
O'Scannlain, Diarmuid	9	1986	66
Trott, Stephen	9	1988	64
Rymer, Pamela Ann	9	1989	62
Nelson, Thomas G.	9	1990	67
Kleinfeld, Andrew J.	9	1991	58
Hawkins, Michael Daly	9	1994	58
Tashima, A. Wallace	9	1996	69
Thomas, Sidney R.	9	1996	50
Seymour, Stephanie	10	1979	63
Tacha, Deanell Reece	10	1985	57
Ebel, David M.	10	1987	63

Table A Continued
Federal Circuit Court Judges in the Tournament

Judge	(A) Circuit	(B) Year Appointed	(C) Age in 2003
Kelly, Paul J.	10	1992	63
Henry, Robert H.	10	1994	50
Murphy, Michael R.	10	1995	56
Briscoe, Mary Beck	10	1995	56
Lucero, Carlos F.	10	1995	63
Gerald Bard Tjoflat	11	1975	74
Robert L. Anderson III	11	1979	67
James L. Edmondson	11	1986	56
Stanley F. Birch Jr	11	1990	58
Joel Fredrick Dubina	11	1990	56
Edward Earl Carnes	11	1992	53
Susan Harrell Black	11	1992	60
Rosemary Barkett	11	1994	64
Frank M. Hull	11	1997	55
Stanley Marcus	11	1997	57
Edwards, Harry	DC	1980	63
Ginsburg, Douglas H.	DC	1986	57
Sentelle, David Bryan	DC	1987	60
Randolph, Arthur	DC	1990	60
Henderson, Karen	DC	1990	59
Tatel, David S.	DC	1994	61
Rogers, Judith Wilson	DC	1994	64
Garland, Merrick B.	DC	1997	51

Age in 2003 is calculated as 2003 subtracted by the year of the judge's birth.

Table B

Published Opinions Written from 1998-2000

Judge	(A) Total Number of Published Opinions	(B) Total Number of Published Majority Opinions	(C) Circuit
Posner, Richard	269	254	7
Easterbrook, Frank	233	213	7
Flaum, Joel	202	192	7
Wood, Diane	194	173	7
Ripple, Kenneth	182	151	7
Kanne, Michael	177	176	7
Morris S. Arnold	175	152	8
Coffey, John	168	162	7
James B. Loken	167	147	8
Roger L. Wollman	158	154	8
Evans, Terrence	153	135	7
Rovner, Ilana	152	122	7
Bruce M. Selya	149	145	1
Reinhardt, Stephen	142	94	9
Pasco M. Bowman	140	137	8
O'Scannlain, Diarmuid	139	94	9
Juan R. Torruella	138	128	1
Boudin, Michael	135	131	1
Smith, Jerry	132	118	5
Karen Nelson Moore	130	94	6
Ronald Lee Gilman	124	99	6
Manion, Daniel	122	102	7
Sandra L. Lynch	120	113	1
Ebel, David M.	114	102	10
Paul V. Niemeyer	113	92	4
Boggs, Danny	113	89	6
Diana E. Murphy	111	106	8
Jacobs, Dennis	109	92	2
Gerald Bard Tjoflat	108	98	11
DeMoss, Harold R. Jr.	108	78	5
Garza, Emilio	106	75	5
Murphy, Michael R.	106	102	10
Trott, Stephen	105	86	9
Stewart, Carl E.	104	102	5
Wilkinson III, J. Harvie	103	86	4
Calabresi, Guido	101	84	2
Jones, Edith	101	80	5

Table B Continued
Published Opinions Written from 1998-2000

Judge	(A) Total Number of Published Opinions	(B) Total Number of Published Majority Opinions	(C) Circuit
Pregerson, Harry	99	82	9
Tashima, A. Wallace	99	76	9
Tacha, Deanell Reece	99	92	10
Higginbotham, Patrick	97	94	5
Walker, John M.	95	91	2
Benavides, Fortunato	94	83	5
Stanley F. Birch Jr	92	82	11
Sloviter, Dolores K.	92	80	3
Kelly, Paul J.	92	81	10
Dennis, James L.	89	49	5
Rosemary Barkett	88	65	11
Edward Earl Carnes	86	72	11
King, Carolyn Dineen	85	75	5
Thomas, Sidney R.	85	67	9
Briscoe, Mary Beck	85	59	10
Wiener, Jacques	83	75	5
Lucero, Carlos F.	83	70	10
Kleinfeld, Andrew J.	82	52	9
Tatel, David S.	82	65	DC
Jolly, E. Grady	82	72	5
Luttig, J. Michael	81	62	4
Eric L. Clay	81	62	6
Cabranes, Jose	81	71	2
Rogers, Judith	79	73	DC
Ginsburg, Douglas H.	79	77	DC
Hawkins, Michael Daly	78	60	9
Randolph, Arthur	77	61	DC
Sentelle, David Bryan	77	60	DC
Diana Gribbon Motz	76	62	4
Henderson, Karen	76	51	DC
Schroeder, Mary M	73	68	9
Edwards, Harry	71	64	DC
H. Emory Widener, Jr.	71	51	4
Barksdale, Rhesa	71	60	5
Alito, Samuel A.	70	52	3
Davis, W. Eugene	70	69	5
R. Guy Cole, Jr.	69	55	6

Table B Continued
Published Opinions Written from 1998-2000

Judge	(A) Total Number of Published Opinions	(B) Total Number of Published Majority Opinions	(C) Circuit
Seymour, Stephanie	68	63	10
Kozinski, Alex	67	40	9
Rymer, Pamela Ann	67	45	9
Alice M. Batchelder	67	48	6
Parker, Fred I.	66	58	2
Henry, Robert H.	63	56	10
Susan Harrell Black	59	49	11
Karen J. Williams	59	50	4
James L. Edmondson	58	51	11
William W. Wilkins	58	47	4
Nygaard, Richard L.	57	45	3
Rendell, Marjorie	55	39	3
M. Blane Michael	55	38	4
McKee, Theodore A.	55	41	3
Garland, Merrick B.	55	52	DC
Roth, Jane R.	54	41	3
Frank M. Hull	54	48	11
Robert L. Anderson III	54	44	11
Stanley Marcus	50	49	11
Boyce F. Martin, Jr.	48	41	6
Martha Daughtrey	46	38	6
Nelson, Thomas G.	45	39	9
Joel Fredrick Dubina	44	40	11
Scirica, Anthony J.	38	36	3

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 defined as those who are in the top 50% of judges based on the number of published opinions (majority, concurrences, and dissents) (A). Bottom judges defined as those who are in the bottom 50% of judges based on the number of published opinions (majority, concurrences, and dissents) (A).

Table C

Published Opinions Written from 1998-2000
Adjusted for Inter-Circuit Differences

Judge	(A) Total Number of Published Opinions (adjusted for circuit variation)	Z-Score of (A)	(B) Circuit
Posner, Richard	269.0	3.60**	7
Reinhardt, Stephen	237.1	2.23**	9
O'Scannlain, Diarmuid	234.1	2.10**	9
Easterbrook, Frank	233.0	2.05**	7
Karen Nelson Moore	230.5	1.94*	6
Ronald Lee Gilman	224.5	1.69*	6
Gerald Bard Tjoflat	223.9	1.66*	11
Smith, Jerry	223.2	1.63	5
Paul V. Niemeyer	221.2	1.54	4
Sloviter, Dolores K.	217.1	1.37	3
Boggs, Danny	213.5	1.21	6
Wilkinson III, J. Harvie	211.2	1.12	4
Ebel, David M.	210.5	1.09	10
Morris S. Arnold	210.0	1.06	8
Stanley F. Birch Jr	207.9	0.97	11
Rosemary Barkett	203.9	0.80	11
Jacobs, Dennis	203.8	0.80	2
Murphy, Michael R.	202.5	0.74	10
Flaum, Joel	202.0	0.72	7
James B. Loken	202.0	0.72	8
Edward Earl Carnes	201.9	0.72	11
Trott, Stephen	200.1	0.64	9
DeMoss, Harold R. Jr.	199.2	0.60	5
Bruce M. Selya	198.7	0.58	1
Garza, Emilio	197.2	0.51	5
Calabresi, Guido	195.8	0.45	2
Tacha, Deanell Reece	195.5	0.44	10
Stewart, Carl E.	195.2	0.43	5
Alito, Samuel A.	195.1	0.42	3
Pregerson, Harry	194.1	0.38	9
Tashima, A. Wallace	194.1	0.38	9
Wood, Diane	194.0	0.38	7
Roger L. Wollman	193.0	0.33	8

Table C Continued
Published Opinions Written from 1998-2000
Adjusted for Inter-Circuit Differences

Judge	(A) Total Number of Published Opinions (adjusted for circuit variation)	Z-Score of (A)	(B) Circuit
Tatel, David S.	192.7	0.32	DC
Jones, Edith	192.2	0.30	5
Walker, John M.	189.8	0.20	2
Ginsburg, Douglas H.	189.7	0.19	DC
Rogers, Judith	189.7	0.19	DC
Luttig, J. Michael	189.2	0.17	4
Kelly, Paul J.	188.5	0.14	10
Higginbotham, Patrick E.	188.2	0.13	5
Juan R. Torruella	187.7	0.11	1
Randolph, Arthur	187.7	0.11	DC
Sentelle, David Bryan	187.7	0.11	DC
Henderson, Karen	186.7	0.06	DC
Benavides, Fortunato	185.2	0.00	5
Boudin, Michael	184.7	-0.02	1
Diana Gribbon Motz	184.2	-0.04	4
Nygaard, Richard L.	182.1	-0.13	3
Ripple, Kenneth	182.0	-0.14	7
Edwards, Harry	181.7	-0.15	DC
Eric L. Clay	181.5	-0.16	6
Briscoe, Mary Beck	181.5	-0.16	10
Dennis, James L.	180.2	-0.22	5
Thomas, Sidney R.	180.1	-0.22	9
McKee, Theodore A.	180.1	-0.22	3
Rendell, Marjorie	180.1	-0.22	3
Lucero, Carlos F.	179.5	-0.25	10
H. Emory Widener, Jr.	179.2	-0.26	4
Roth, Jane R.	179.1	-0.26	3
Kleinfeld, Andrew J.	177.1	-0.35	9
Kanne, Michael	177.0	-0.35	7
King, Carolyn Dineen	176.2	-0.39	5
Cabranes, Jose	175.8	-0.40	2
Pasco M. Bowman	175.0	-0.44	8
Susan Harrell Black	174.9	-0.44	11
Wiener, Jacques	174.2	-0.47	5

Table C Continued
Published Opinions Written from 1998-2000
Adjusted for Inter-Circuit Differences

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

** 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 7th 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 defined as those who are in the top 50% of judges based on the number of published opinions (majority, concurrences, and dissents) adjusted for intercircuit differences (A). Bottom judges defined as those who are in the bottom 50% of judges based on the number of published opinions (majority, concurrences, and dissents) adjusted for intercircuit differences (A).

Table D

Citations to Opinions Published from 1998-2000

Judge	(A) Total Outside Circuit Citations	Z-Score of normalized (A)	(B) SCT Citations	Z-Score of normalized (B)	(C) Law Review and Periodical Citations	Z-Score of normalized (C)	(D) Self-Citations	Z-Score of normalized (D)	(E) Circuit
Posner, Richard	1406	2.61**	16	2.31**	1033	2.41**	392	2.35**	7
Easterbrook, Frank	1340	2.52**	14	2.11**	790	1.83*	257	1.95*	7
Sandra L. Lynch	1023	1.99**	5	0.62	684	1.52	178	1.60	1
Bruce M. Selya	949	1.85*	3	-0.04	727	1.65*	364	2.28**	1
Kelly, Paul J.	799	1.51	0	-2.29**	388	0.30	103	1.07	10
Kanne, Michael	768	1.44	4	0.32	512	0.90	181	1.61	7
Flaum, Joel	743	1.37	3	-0.04	613	1.29	126	1.27	7
Ripple, Kenneth	730	1.34	4	0.32	545	1.03	168	1.54	7
Wood, Diane	678	1.20	3	-0.04	513	0.90	127	1.27	7
Wilkinson III, J. Harvie	662	1.15	4	0.32	648	1.41	23	-0.36	4
Edward Earl Carnes	648	1.11	4	0.32	448	0.61	113	1.16	11
Smith, Jerry	637	1.08	3	-0.04	622	1.32	53	0.44	5
Ebel, David M.	634	1.07	4	0.32	503	0.86	123	1.24	10
Paul V. Niemeyer	623	1.03	10	1.60	390	0.31	46	0.31	4
Walker, John M.	616	1.01	4	0.32	667	1.47	36	0.07	2
Reinhardt, Stephen	605	0.98	11	1.75*	788	1.83*	66	0.65	9
Pasco M. Bowman	573	0.87	4	0.32	487	0.79	42	0.22	8
King, Carolyn Dineen	572	0.87	5	0.62	490	0.80	60	0.56	5
Calabresi, Guido	566	0.85	7	1.09	604	1.25	55	0.48	2
Jolly, E. Grady	554	0.80	2	-0.51	422	0.48	20	-0.49	5
Karen J. Williams	552	0.80	8	1.28	437	0.56	22	-0.40	4
Juan R. Torruella	544	0.77	3	-0.04	444	0.59	84	0.88	1
Roger L. Wollman	541	0.76	5	0.62	414	0.44	246	1.90	8
Frank M. Hull	525	0.70	6	0.87	241	-0.73	21	-0.44	11
Scirica, Anthony J.	523	0.69	2	-0.51	236	-0.77	19	-0.54	3
James B. Loken	518	0.67	3	-0.04	446	0.60	61	0.57	8

Table D Continued

Citations to Opinions Published from 1998-2000

Judge	(A) Total Outside Circuit Citations	Z-Score of normalized (A)	(B) SCT Citations	Z-Score of normalized (B)	(C) Law Review and Periodical Citations	Z-Score of normalized (C)	(D) Self-Citations	Z-Score of normalized (D)	(E) Circuit
Karen Nelson Moore	517	0.67	4	0.32	403	0.38	160	1.49	6
Boggs, Danny	507	0.63	6	0.87	424	0.49	39	0.15	6
Gerald Bard Tjoflat	507	0.63	3	-0.04	498	0.84	65	0.63	11
Jacobs, Dennis	503	0.62	4	0.32	450	0.62	36	0.07	2
Benavides, Fortunato	469	0.48	2	-0.51	456	0.65	22	-0.40	5
Manion, Daniel	467	0.47	4	0.32	386	0.29	80	0.83	7
Murphy, Michael R.	454	0.42	6	0.87	368	0.19	109	1.13	10
Coffey, John	446	0.38	2	-0.51	465	0.69	173	1.57	7
Boudin, Michael	443	0.37	6	0.87	346	0.05	82	0.86	1
Tacha, Deanell Reece	442	0.37	2	-0.51	509	0.89	83	0.87	10
Diana Gribbon Motz	430	0.31	4	0.32	253	-0.62	18	-0.59	4
Eric L. Clay	418	0.26	3	-0.04	383	0.27	113	1.16	6
Stanley F. Birch Jr	417	0.25	7	1.09	401	0.37	25	-0.28	11
DeMoss, Harold R. Jr.	412	0.23	6	0.87	293	-0.30	34	0.02	5
Ronald Lee Gilman	409	0.22	4	0.32	400	0.37	75	0.77	6
Garza, Emilio	406	0.20	6	0.87	385	0.28	22	-0.40	5
Evans, Terrence	405	0.20	4	0.32	413	0.44	31	-0.07	7
Sloviter, Dolores K.	403	0.19	5	0.62	394	0.33	27	-0.20	3
Cabranes, Jose	396	0.15	1	-1.16	354	0.10	16	-0.70	2
Nygaard, Richard L.	394	0.14	1	-1.16	191	-1.23	11	-1.06	3
Diana E. Murphy	393	0.14	0	-2.29**	386	0.29	26	-0.24	8
Rovner, Ilana	392	0.13	3	-0.04	376	0.23	66	0.65	7
O'Scannlain, Diarmuid	386	0.10	6	0.87	741	1.69	29	-0.13	9
Jones, Edith	380	0.07	3	-0.04	589	1.20	25	-0.28	5
Luttig, J. Michael	378	0.06	8	1.28	476	0.74	16	-0.70	4
Morris S. Arnold	369	0.02	6	0.87	357	0.12	60	0.56	8

Table D Continued

Citations to Opinions Published from 1998-2000

Judge	(A) Total Outside Circuit Citations	Z-Score of normalized (A)	(B) SCT Citations	Z-Score of normalized (B)	(C) Law Review and Periodical Citations	Z-Score of normalized (C)	(D) Self-Citations	Z-Score of normalized (D)	(E) Circuit
Barksdale, Rhesa	368	0.01	5	0.62	334	-0.02	26	-0.24	5
Tashima, A. Wallace	367	0.01	4	0.32	317	-0.13	34	0.02	9
Higginbotham, Patrick	364	-0.01	5	0.62	544	1.03	29	-0.13	5
Wiener, Jacques	349	-0.09	3	-0.04	401	0.37	21	-0.44	5
Lucero, Carlos F.	346	-0.11	6	0.87	305	-0.22	72	0.73	10
Tatel, David S.	345	-0.11	7	1.09	212	-1.00	40	0.17	DC
William W. Wilkins	341	-0.14	5	0.62	374	0.22	29	-0.13	4
Seymour, Stephanie	340	-0.14	5	0.62	306	-0.21	43	0.24	10
Stewart, Carl E.	319	-0.27	4	0.32	527	0.96	36	0.07	5
James L. Edmondson	317	-0.28	7	1.09	390	0.31	16	-0.70	11
Davis, W. Eugene	309	-0.33	2	-0.51	440	0.57	9	-1.25	5
Thomas, Sidney R.	306	-0.35	6	0.87	314	-0.16	34	0.02	9
R. Guy Cole, Jr.	293	-0.43	6	0.87	212	-1.00	56	0.49	6
Briscoe, Mary Beck	284	-0.49	6	0.87	247	-0.67	30	-0.10	10
Alice M. Batchelder	281	-0.51	4	0.32	201	-1.12	16	-0.70	6
Trott, Stephen	281	-0.51	4	0.32	422	0.48	26	-0.24	9
Pregerson, Harry	278	-0.53	2	-0.51	241	-0.73	23	-0.36	9
Rosemary Barkett	276	-0.55	4	0.32	251	-0.64	15	-0.76	11
Henry, Robert H.	276	-0.55	4	0.32	232	-0.81	45	0.28	10
Alito, Samuel A.	263	-0.64	4	0.32	240	-0.73	12	-0.98	3
Parker, Fred I.	261	-0.66	3	-0.04	285	-0.36	14	-0.83	2
M. Blane Michael	248	-0.76	3	-0.04	200	-1.13	17	-0.64	4
Robert L. Anderson III	246	-0.77	1	-1.16	273	-0.46	11	-1.06	11
McKee, Theodore A.	244	-0.79	1	-1.16	172	-1.45	22	-0.40	3
Rendell, Marjorie	244	-0.79	0	-2.29**	215	-0.97	17	-0.64	3
Stanley Marcus	237	-0.84	1	-1.16	334	-0.02	51	0.40	11

Table D Continued

Citations to Opinions Published from 1998-2000

Judge	(A) Total Outside Circuit Citations	Z-Score of normalized (A)	(B) SCT Citations	Z-Score of normalized (B)	(C) Law Review and Periodical Citations	Z-Score of normalized (C)	(D) Self-Citations	Z-Score of normalized (D)	(E) Circuit
Roth, Jane R.	232	-0.89	2	-0.51	199	-1.14	13	-0.90	3
Schroeder, Mary M	230	-0.90	1	-1.16	237	-0.76	3	-2.30**	9
Martha Daughtrey	224	-0.95	2	-0.51	160	-1.61	2	-2.68**	6
Susan Harrell Black	220	-0.99	1	-1.16	185	-1.30	8	-1.36	11
Randolph, Arthur	216	-1.02	2	-0.51	292	-0.31	36	0.07	DC
Joel Fredrick Dubina	214	-1.04	4	0.32	152	-1.72*	8	-1.36	11
Rogers, Judith	214	-1.04	2	-0.51	193	-1.20	24	-0.31	DC
Hawkins, Michael Daly	214	-1.04	4	0.32	196	-1.17	11	-1.06	9
Dennis, James L.	202	-1.15	0	-2.29**	204	-1.08	18	-0.59	5
Ginsburg, Douglas H.	193	-1.24	2	-0.51	263	-0.54	32	-0.04	DC
Edwards, Harry	177	-1.41	1	-1.16	229	-0.84	22	-0.40	DC
Garland, Merrick B.	169	-1.50	1	-1.16	176	-1.40	72	0.73	DC
H. Emory Widener, Jr.	169	-1.50	1	-1.16	305	-0.22	4	-2.02**	4
Nelson, Thomas G.	167	-1.52	2	-0.51	117	-2.28	2	-2.68**	9
Kleinfeld, Andrew J.	157	-1.64	0	-2.29**	164	-1.56	19	-0.54	9
Kozinski, Alex	149	-1.74*	4	0.32	216	-0.96	10	-1.15	9
Sentelle, David Bryan	148	-1.76*	2	-0.51	148	-1.78*	26	-0.24	DC
Boyce F. Martin, Jr.	124	-2.10**	0	-2.29**	138	-1.93*	17	-0.64	6
Rymer, Pamela Ann	112	-2.30**	0	-2.29**	206	-1.06	3	-2.30**	9
Henderson, Karen	109	-2.35**	0	-2.29**	116	-2.30**	35	0.04	DC

** 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 measured to 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 defined as those who are in the top 50% of judges based on the number of outside circuit citations (A). Bottom judges defined as those who are in the bottom 50% 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 from 1998-2000

Judge	(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 L. Lynch	734	2.56**	9.03	1.73*	1
Easterbrook, Frank	667	2.33**	6.25	0.73	7
Kelly, Paul J.	654	2.28**	9.85	1.97**	10
Posner, Richard	570	1.95*	5.49	0.37	7
Bruce M. Selya	516	1.71*	6.50	0.83	1
Scirica, Anthony J.	496	1.61	14.50	3.04**	3
Frank M. Hull	460	1.43	10.90	2.25**	11
Karen J. Williams	455	1.40	11.02	2.28**	4
Edward Earl Carnes	444	1.34	8.92	1.70*	11
Wilkinson III, J. Harvie	425	1.24	7.64	1.28	4
Walker, John M.	423	1.22	6.75	0.93	2
Jolly, E. Grady	415	1.18	7.65	1.28	5
Ebel, David M.	412	1.16	6.22	0.71	10
King, Carolyn Dineen	407	1.13	7.61	1.27	5
Smith, Jerry	389	1.02	5.30	0.27	5
Reinhardt, Stephen	374	0.93	6.33	0.76	9
Calabresi, Guido	371	0.91	6.67	0.90	2
Pasco M. Bowman	371	0.91	4.16	-0.39	8
Paul V. Niemeyer	365	0.87	6.77	0.94	4
Jacobs, Dennis	345	0.73	5.40	0.32	2
Flaum, Joel	342	0.71	3.86	-0.60	7
Benavides, Fortunato	341	0.70	5.64	0.44	5
Ripple, Kenneth	340	0.69	4.82	0.01	7
Nygaard, Richard L.	336	0.67	8.76	1.65*	3
Kanne, Michael	328	0.61	4.36	-0.27	7

Table E Continued

Outside Circuit Citations to Opinions Controlling for Total Number of Opinions Published from 1998-2000

Judge	(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
Wood, Diane	327	0.60	3.87	-0.60	7
Luttig, J. Michael	322	0.56	6.05	0.63	4
Boggs, Danny	320	0.55	5.63	0.44	6
Karen Nelson Moore	315	0.51	5.45	0.35	6
Gerald Bard Tjoflat	313	0.49	5.11	0.17	11
Barksdale, Rhesa	309	0.46	6.13	0.67	5
Eric L. Clay	309	0.46	6.65	0.89	6
Diana Gribbon Motz	307	0.45	6.87	0.98	4
William W. Wilkins	295	0.35	7.19	1.11	4
Stanley F. Birch Jr	292	0.33	5.05	0.14	11
DeMoss, Harold R. Jr.	288	0.29	5.24	0.24	5
O'Scannlain, Diarmuid F.	284	0.26	4.07	-0.45	9
Juan R. Torruella	281	0.23	4.20	-0.37	1
Sloviter, Dolores K.	277	0.20	5.04	0.13	3
Tacha, Deanell Reece	277	0.20	4.79	-0.01	10
Ronald Lee Gilman	274	0.17	4.05	-0.47	6
Roger L. Wollman	274	0.17	3.51	-0.86	8
Garza, Emilio	272	0.15	5.35	0.30	5
Wiener, Jacques Loeb Jr.	267	0.11	4.64	-0.09	5
Manion, Daniel	265	0.09	4.56	-0.14	7
Tatel, David S.	265	0.09	5.11	0.17	DC
Jones, Edith	262	0.06	4.74	-0.04	5
James B. Loken	258	0.03	3.51	-0.86	8
Tashima, A. Wallace	257	0.02	4.79	-0.01	9
Seymour, Stephanie	256	0.01	5.37	0.30	10

Table E Continued

Outside Circuit Citations to Opinions Controlling for Total Number of Opinions Published from 1998-2000

Judge	(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
James Larry Edmondson	255	0.00	6.22	0.71	11
Murphy, Michael R.	253	-0.02	4.45	-0.21	10
Davis, W. Eugene	250	-0.05	4.46	-0.20	5
Alice M. Batchelder	250	-0.05	5.83	0.53	6
Diana E. Murphy	245	-0.10	3.69	-0.73	8
Lucero, Carlos F.	240	-0.15	4.89	0.05	10
Thomas, Sidney R.	233	-0.22	4.51	-0.17	9
Higginbotham, Patrick E.	232	-0.23	3.81	-0.64	5
Rovner, Ilana	232	-0.23	3.18	-1.13	7
Evans, Terrence	230	-0.25	3.00	-1.29	7
Cabranes, Jose	229	-0.26	5.58	0.41	2
Boudin, Michael	223	-0.33	3.31	-1.03	1
Coffey, John	222	-0.34	2.75	-1.54	7
McKee, Theodore A.	218	-0.38	5.90	0.57	3
R. Guy Cole, Jr.	218	-0.38	5.29	0.27	6
Robert L. Anderson III	218	-0.38	5.55	0.40	11
M. Blane Michael	217	-0.39	6.45	0.81	4
Martha Craig Daughtrey	212	-0.45	5.89	0.56	6
Briscoe, Mary Beck	207	-0.51	4.76	-0.02	10
Alito, Samuel A.	205	-0.53	4.94	0.08	3
Rosemary Barkett	205	-0.53	4.17	-0.39	11
Henry, Robert H.	201	-0.58	4.91	0.06	10
Trott, Stephen	198	-0.62	3.23	-1.09	9
Rendell, Marjorie	196	-0.64	6.21	0.70	3
Roth, Jane R.	196	-0.64	5.54	0.39	3

Table E Continued

Outside Circuit Citations to Opinions Controlling for Total Number of Opinions Published from 1998-2000

Judge	(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
Parker, Fred I.	193	-0.68	4.50	-0.18	2
Stanley Marcus	190	-0.72	4.84	0.02	11
Morris S. Arnold	188	-0.74	2.36	-1.95*	8
Schroeder, Mary M	187	-0.76	3.38	-0.96	9
Joel Fredrick Dubina	187	-0.76	5.35	0.30	11
Susan Harrell Black	172	-0.96	4.47	-0.20	11
Rogers, Judith	167	-1.03	2.86	-1.42	DC
Pregerson, Harry	166	-1.04	3.39	-0.96	9
Randolph, Arthur	164	-1.07	3.48	-0.89	DC
Stewart, Carl E.	163	-1.09	3.13	-1.18	5
Hawkins, Michael Daly	163	-1.09	3.55	-0.83	9
Dennis, James L.	156	-1.20	3.78	-0.66	5
Ginsburg, Douglas H.	149	-1.31	2.49	-1.80*	DC
Nelson, Thomas G.	148	-1.32	4.18	-0.38	9
Edwards, Harry	145	-1.37	2.72	-1.56	DC
Garland, Merrick B.	143	-1.41	3.25	-1.07	DC
H. Emory Widener, Jr.	139	-1.48	3.29	-1.04	4
Kozinski, Alex	137	-1.51	3.65	-0.75	9
Kleinfeld, Andrew J.	133	-1.58	2.96	-1.33	9
Sentelle, David Bryan	113	-1.98**	2.45	-1.85*	DC
Boyce F. Martin, Jr.	110	-2.04**	2.98	-1.32	6
Henderson, Karen	98	-2.32**	2.08	-2.30**	DC
Rymer, Pamela Ann	96	-2.37**	2.47	-1.83*	9

** 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 measured to 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 20 citation-receiving cases (A): $\chi^2 = 15.466$ (11 d.f.) ($p \leq 0.169$). Top judges defined as those who are in the top 50% of judges based on the number of outside circuit citations to the top 20 citation-receiving cases (A). Bottom judges defined as those who are in the bottom 50% of judges based on the number of outside circuit citations to the top 20 citation-receiving cases (A).

Table F

Invocations to Opinions Published from 1998-2000

Judge	(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
Posner, Richard	176	3.90**	0.65	6.68**	97.7%	7
Easterbrook, Frank	103	3.36**	0.44	4.67**	99.0%	7
Calabresi, Guido	23	1.85*	0.23	2.35**	91.3%	2
Wilkinson III, J. Harvie	19	1.66*	0.18	1.73*	73.7%	4
Boudin, Michael	13	1.30	0.10	0.70	84.6%	1
Higginbotham, Patrick E.	12	1.22	0.12	0.96	41.7%	5
O'Scannlain, Diarmuid F.	11	1.14	0.08	0.43	54.5%	9
Jones, Edith	11	1.14	0.11	0.83	81.8%	5
Wood, Diane	10	1.05	0.05	0.01	20.0%	7
Luttig, J. Michael	10	1.05	0.12	0.96	80.0%	4
Ripple, Kenneth	9	0.95	0.05	0.01	55.6%	7
Walker, John M.	9	0.95	0.09	0.56	77.8%	2
Gerald Bard Tjoflat	9	0.95	0.08	0.43	100.0%	11
Morris S. Arnold	8	0.84	0.05	0.01	12.5%	8
Bruce M. Selya	8	0.84	0.05	0.01	75.0%	1
James Larry Edmondson	8	0.84	0.14	1.23	100.0%	11
Rovner, Ilana	8	0.84	0.05	0.01	12.5%	7
Ronald Lee Gilman	8	0.84	0.06	0.15	25.0%	6
Sandra L. Lynch	7	0.72	0.06	0.15	71.4%	1
King, Carolyn Dineen	7	0.72	0.08	0.43	71.4%	5
Kleinfeld, Andrew J.	7	0.72	0.09	0.56	42.9%	9
Paul V. Niemeyer	7	0.72	0.06	0.15	71.4%	4
Reinhardt, Stephen	7	0.72	0.05	0.01	0.0%	9
Eric L. Clay	7	0.72	0.09	0.56	71.4%	6
Cabranes, Jose	7	0.72	0.09	0.56	71.4%	2
Evans, Terrence	6	0.58	0.04	-0.13	100.0%	7

Table F Continued

Invocations to Opinions Published from 1998-2000

Judge	(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
Karen Nelson Moore	6	0.58	0.05	0.01	16.7%	6
Alito, Samuel A.	5	0.42	0.07	0.29	40.0%	3
Manion, Daniel	5	0.42	0.04	-0.13	40.0%	7
Trott, Stephen	5	0.42	0.05	0.01	60.0%	9
Boggs, Danny	5	0.42	0.04	-0.13	60.0%	6
Kozinski, Alex	5	0.42	0.07	0.29	100.0%	9
Jacobs, Dennis	5	0.42	0.05	0.01	0.0%	2
Randolph, Arthur Raymond	5	0.42	0.06	0.15	60.0%	DC
Rogers, Judith Ann Wilson	5	0.42	0.06	0.15	80.0%	DC
Benavides, Fortunato Pedro	5	0.42	0.05	0.01	60.0%	5
Garza, Emilio	5	0.42	0.05	0.01	0.0%	5
Tatel, David S.	5	0.42	0.06	0.15	60.0%	DC
James B. Loken	5	0.42	0.03	-0.27	20.0%	8
Thomas, Sidney R.	4	0.24	0.05	0.01	25.0%	9
Roth, Jane R.	4	0.24	0.07	0.29	25.0%	3
Seymour, Stephanie	4	0.24	0.06	0.15	75.0%	10
Dennis, James L.	4	0.24	0.04	-0.13	25.0%	5
Diana Gribbon Motz	4	0.24	0.05	0.01	50.0%	4
R. Guy Cole, Jr.	4	0.24	0.06	0.15	100.0%	6
Stanley F. Birch Jr	4	0.24	0.04	-0.13	50.0%	11
Edward Earl Carnes	3	0.01	0.03	-0.27	100.0%	11
Flaum, Joel	3	0.01	0.01	-0.56	66.7%	7
Rendell, Marjorie	3	0.01	0.05	0.01	33.3%	3
Rosemary Barkett	3	0.01	0.03	-0.27	66.7%	11
Smith, Jerry	3	0.01	0.02	-0.41	0.0%	5
Nelson, Thomas G.	3	0.01	0.07	0.29	0.0%	9

Table F Continued

Invocations to Opinions Published from 1998-2000

Judge	(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
Pregerson, Harry	3	0.01	0.03	-0.27	66.7%	9
M. Blane Michael	3	0.01	0.05	0.01	66.7%	4
Kanne, Michael	3	0.01	0.02	-0.41	66.7%	7
Scirica, Anthony J.	2	-0.29	0.05	0.01	100.0%	3
Schroeder, Mary M	2	-0.29	0.03	-0.27	50.0%	9
McKee, Theodore A.	2	-0.29	0.04	-0.13	0.0%	3
William W. Wilkins	2	-0.29	0.03	-0.27	50.0%	4
Tashima, A. Wallace	2	-0.29	0.02	-0.41	0.0%	9
Coffey, John	2	-0.29	0.01	-0.56	50.0%	7
Edwards, Harry	2	-0.29	0.03	-0.27	50.0%	DC
Wiener, Jacques Loeb Jr.	2	-0.29	0.02	-0.41	0.0%	5
Susan Harrell Black	2	-0.29	0.03	-0.27	0.0%	11
H. Emory Widener, Jr.	2	-0.29	0.03	-0.27	0.0%	4
Karen J. Williams	2	-0.29	0.03	-0.27	50.0%	4
Frank M. Hull	2	-0.29	0.04	-0.13	50.0%	11
Lucero, Carlos F.	2	-0.29	0.02	-0.41	0.0%	10
Ebel, David M.	1	-0.71	0.01	-0.56	100.0%	10
Jolly, E. Grady	1	-0.71	0.01	-0.56	0.0%	5
Juan R. Torruella	1	-0.71	0.01	-0.56	0.0%	1
DeMoss, Harold R. Jr.	1	-0.71	0.01	-0.56	0.0%	5
Ginsburg, Douglas H.	1	-0.71	0.01	-0.56	0.0%	DC
Sloviter, Dolores K.	1	-0.71	0.01	-0.56	100.0%	3
Sentelle, David Bryan	1	-0.71	0.01	-0.56	100.0%	DC
Rymer, Pamela Ann	1	-0.71	0.01	-0.56	0.0%	9
Briscoe, Mary Beck	1	-0.71	0.01	-0.56	0.0%	10
Alice M. Batchelder	1	-0.71	0.01	-0.56	100.0%	6

Table F Continued

Invocations to Opinions Published from 1998-2000

Judge	(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
Pasco M. Bowman	1	-0.71	0.01	-0.56	100.0%	8
Roger L. Wollman	1	-0.71	0.01	-0.56	100.0%	8
Robert Lanier Anderson III	0	-1.42	0.00	-0.71	–	11
Stewart, Carl E.	0	-1.42	0.00	-0.71	–	5
Garland, Merrick B.	0	-1.42	0.00	-0.71	–	DC
Nygaard, Richard L.	0	-1.42	0.00	-0.71	–	3
Davis, W. Eugene	0	-1.42	0.00	-0.71	–	5
Joel Fredrick Dubina	0	-1.42	0.00	-0.71	–	11
Kelly, Paul J.	0	-1.42	0.00	-0.71	–	10
Parker, Fred I.	0	-1.42	0.00	-0.71	–	2
Murphy, Michael R.	0	-1.42	0.00	-0.71	–	10
Martha Craig Daughtrey	0	-1.42	0.00	-0.71	–	6
Henry, Robert H.	0	-1.42	0.00	-0.71	–	10
Henderson, Karen LeCraft	0	-1.42	0.00	-0.71	–	DC
Diana E. Murphy	0	-1.42	0.00	-0.71	–	8
Hawkins, Michael Daly	0	-1.42	0.00	-0.71	–	9
Boyce F. Martin, Jr.	0	-1.42	0.00	-0.71	–	6
Barksdale, Rhesa Hawkins	0	-1.42	0.00	-0.71	–	5
Tacha, Deanell Reece	0	-1.42	0.00	-0.71	–	10
Stanley Marcus	0	-1.42	0.00	-0.71	–	11

** 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 defined as those who are in the top 50% of judges based on the number of invocations (A). Bottom judges defined as those who are in the bottom 50% of judges based on the number of invocations (A).

Table G

Number of Opposing Opinions and Independence Ratings for the period 1998-2000

Judge	(A) Independence Rating	(B) Number of Dissents and Concurrences	(C) Adjusted Dissents and Concurrences for Intercircuit Differences	Z-Score for (C)	(D) Circuit
Ebel, David M.	0.000	12	24.5	0.16	10
Robert L. Anderson III	0.000	10	23.7	0.06	11
Luttig, J. Michael	-0.006	19	26.2	0.35	4
Alito, Samuel A.	-0.015	18	28.7	0.65	3
Stewart, Carl E.	-0.017	2	10.4	-1.49	5
Jolly, E. Grady	-0.018	10	18.4	-0.56	5
Juan R. Torruella	-0.018	10	26.9	0.44	1
Wood, Diane	-0.018	21	27.0	0.44	7
Posner, Richard	0.019	15	21.0	-0.26	7
Edward Earl Carnes	-0.022	14	27.7	0.53	11
Scirica, Anthony J.	-0.023	2	12.7	-1.22	3
Schroeder, Mary M	0.023	5	5.0	-2.12**	9
Evans, Terrence	0.024	18	24.0	0.09	7
DeMoss, Harold R. Jr.	0.024	30	38.4	1.78	5
Manion, Daniel	0.028	20	26.0	0.33	7
Garland, Merrick B.	-0.037	3	14.5	-1.01	DC
Easterbrook, Frank	-0.042	20	26.0	0.33	7
Sandra L. Lynch	-0.043	7	23.9	0.09	1
Trott, Stephen	-0.044	19	19.0	-0.49	9
Flaum, Joel	-0.044	10	16.0	-0.84	7
Jones, Edith	-0.045	21	29.4	0.73	5
Rendell, Marjorie	-0.049	16	26.7	0.42	3
Nygaard, Richard L.	0.049	12	22.7	-0.05	3
Ginsburg, Douglas H.	0.052	2	13.5	-1.12	DC
Thomas, Sidney R.	-0.053	18	18.0	-0.60	9
Rosemary Barkett	0.056	23	36.7	1.58	11
Morris S. Arnold	-0.060	23	35.2	1.40	8
McKee, Theodore A.	-0.062	14	24.7	0.18	3
William W. Wilkins	0.065	11	18.2	-0.58	4

Table G Continued

Number of Opposing Opinions and Independence Ratings for the period 1998-2000

Judge	(A) Independence Rating	(B) Number of Dissents and Concurrences	(C) Adjusted Dissents and Concurrences for Intercircuit Differences	Z-Score for (C)	(D) Circuit
King, Carolyn Dineen	-0.067	10	18.4	-0.56	5
Boudin, Michael	0.071	4	20.9	-0.26	1
Davis, W. Eugene	-0.071	1	9.4	-1.61	5
Bruce M. Selya	-0.071	4	20.9	-0.26	1
Sloviter, Dolores K.	-0.077	12	22.7	-0.05	3
Boggs, Danny	-0.082	24	28.2	0.58	6
Kleinfeld, Andrew J.	0.087	30	30.0	0.80	9
Smith, Jerry	-0.093	14	22.4	-0.09	5
James L. Edmondson	0.100	7	20.7	-0.29	11
Joel Fredrick Dubina	0.100	4	17.7	-0.64	11
Wilkinson III, J. Harvie	-0.102	17	24.2	0.12	4
Paul V. Niemeyer	-0.102	21	28.2	0.58	4
Higginbotham, Patrick	-0.036	3	11.4	-1.38	5
Ripple, Kenneth	-0.107	31	37.0	1.61	7
Roth, Jane R.	0.112	13	23.7	0.07	3
Rovner, Ilana	0.112	30	36.0	1.50	7
Kelly, Paul J.	0.115	11	23.5	0.04	10
Tashima, A. Wallace	-0.122	23	23.0	-0.02	9
Parker, Fred I.	-0.124	8	20.0	-0.37	2
Murphy, Michael R.	-0.127	4	16.5	-0.77	10
Reinhardt, Stephen	-0.128	48	48.0	2.90	9
Kozinski, Alex	0.135	27	27.0	0.45	9
Nelson, Thomas G.	0.135	6	6.0	-2.01*	9
Coffey, John	0.139	6	12.0	-1.31	7
Seymour, Stephanie	-0.150	5	17.5	-0.66	10
Dennis, James L.	-0.157	40	48.4	2.95**	5
O'Scannlain, Diarmuid	-0.165	45	45.0	2.55**	9
Sentelle, David Bryan	-0.166	17	28.5	0.63	DC
Edwards, Harry	-0.170	7	18.5	-0.54	DC

Table G Continued

Number of Opposing Opinions and Independence Ratings for the period 1998-2000

Judge	(A) Independence Rating	(B) Number of Dissents and Concurrences	(C) Adjusted Dissents and Concurrences for Intercircuit Differences	Z-Score for (C)	(D) Circuit
Rymer, Pamela Ann	-0.180	22	22.0	-0.14	9
Martha C. Daughtrey	-0.181	8	12.2	-1.29	6
Diana Gribbon Motz	-0.185	14	21.2	-0.23	4
Henry, Robert H.	-0.200	7	19.5	-0.42	10
Jacobs, Dennis	0.203	17	29.0	0.68	2
Wiener, Jacques	0.208	8	16.4	-0.79	5
Randolph, Arthur	0.208	16	27.5	0.51	DC
Briscoe, Mary Beck	-0.210	26	38.5	1.80*	10
Eric L. Clay	-0.212	19	23.2	0.00	6
Henderson, Karen	-0.219	25	36.5	1.56	DC
Rogers, Judith	-0.227	6	17.5	-0.66	DC
Diana E. Murphy	0.228	5	17.2	-0.70	8
Walker, John M.	0.228	4	16.0	-0.84	2
Ronald Lee Gilman	-0.229	25	29.2	0.70	6
Pregerson, Harry	-0.232	17	17.0	-0.72	9
Karen Nelson Moore	-0.233	36	40.2	1.99**	6
Susan Harrell Black	-0.234	10	23.7	0.06	11
Alice M. Batchelder	0.057	19	23.2	0.00	6
Benavides, Fortunato	-0.246	11	19.4	-0.44	5
Garza, Emilio	-0.254	31	39.4	1.90*	5
H. Emory Widener, Jr.	0.257	20	27.2	0.47	4
Karen J. Williams	-0.259	9	16.2	-0.82	4
Hawkins, Michael Daly	-0.261	18	18.0	-0.60	9
Boyce F. Martin, Jr.	0.273	7	11.2	-1.40	6
Barksdale, Rhesa	-0.279	11	19.4	-0.44	5
R. Guy Cole, Jr.	-0.282	14	18.2	-0.58	6
Pasco M. Bowman	-0.283	3	15.2	-0.93	8
Tatel, David S.	-0.287	17	28.5	0.63	DC
Frank M. Hull	0.292	6	19.7	-0.41	11

Table G Continued

Number of Opposing Opinions and Independence Ratings for the period 1998-2000

Judge	(A) Independence Rating	(B) Number of Dissents and Concurrences	(C) Adjusted Dissents and Concurrences for Intercircuit Differences	Z-Score for (C)	(D) Circuit
Gerald Bard Tjoflat	-0.294	10	23.7	0.06	11
Calabresi, Guido	-0.307	17	29.0	0.68	2
Cabranes, Jose	-0.307	10	22.0	-0.14	2
M. Blane Michael	-0.324	17	24.2	0.12	4
Stanley F. Birch Jr	0.333	10	23.7	0.06	11
Tacha, Deanell Reece	-0.389	7	19.5	-0.42	10
Lucero, Carlos F.	-0.400	13	25.5	0.28	10
Kanne, Michael	-0.417	1	7.0	-1.89*	7
Roger L. Wollman	-0.449	4	16.2	-0.82	8
James B. Loken	-0.496	20	32.2	1.05	8
Stanley Marcus	0.542	1	14.7	-0.99	11

** 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 defined as those who are in the top 50% of judges based on the independence rating (A). Bottom judges defined as those who are in the bottom 50% of judges based on the independence rating (A).

Table H

Composite Ranking of Judges With Equal Weighting of Quality, Productivity, and Independence

(Includes only Active Circuit Court Judges 65 or Less in 2003)

Judges (Ranked highest to lowest scoring) based on equal weighting of Quality, Productivity, and Independence	(A) Rank based on Equal Weighted Composite Measure (Best = 1)	(B) Z-Score based on Equal Weighted Composite Measure	(C) Circuit
Posner, Richard	1	3.77**	7
Easterbrook, Frank	2	2.93**	7
Wilkinson III, J. Harvie	3	1.51	4
Paul V. Niemeyer	4	1.34	4
Smith, Jerry	5	1.26	5
Ebel, David M.	6	1.21	10
Edward Earl Carnes	7	1.20	11
Wood, Diane	8	1.14	7
Boggs, Danny	9	1.10	6
Luttig, J. Michael	10	1.08	4
Sandra L. Lynch	11	1.02	1
Karen Nelson Moore	12	0.89	6
Jones, Edith	13	0.88	5
Morris S. Arnold	14	0.87	8
Ronald Lee Gilman	15	0.77	6
Alito, Samuel A.	16	0.69	3
King, Carolyn Dineen	17	0.66	5
Trott, Stephen	18	0.66	9
Rosemary Barkett	19	0.59	11
Boudin, Michael	20	0.57	1
Ripple, Kenneth	21	0.56	7
Higginbotham, Patrick	22	0.52	5
Jacobs, Dennis	23	0.50	2
Kelly, Paul J.	24	0.45	10
Scirica, Anthony J.	25	0.42	3
Walker, John M.	26	0.41	2
Thomas, Sidney R.	27	0.28	9
James L. Edmondson	28	0.21	11
Nygaard, Richard L.	29	0.13	3
Rendell, Marjorie	30	0.12	3
McKee, Theodore A.	31	0.06	3
Diana Gribbon Motz	32	0.05	4
Murphy, Michael R.	33	0.04	10

Table H Continued

Composite Ranking of Judges With Equal Weighting of Quality, Productivity, and Independence

(Includes only Active Circuit Court Judges 65 or Less in 2003)

Judges (Ranked highest to lowest scoring) based on equal weighting of Quality, Productivity, and Independence	(A) Rank based on Equal Weighted Composite Measure (Best = 1)	(B) Z-Score based on Equal Weighted Composite Measure	(C) Circuit
Eric L. Clay	34	0.03	6
Stewart, Carl E.	35	0.02	5
Garza, Emilio	36	0.01	5
William W. Wilkins	37	-0.02	4
Benavides, Fortunato	38	-0.04	5
Evans, Terrence	39	-0.05	7
Stanley F. Birch Jr	40	-0.08	11
Ginsburg, Douglas H.	41	-0.08	DC
Schroeder, Mary M	42	-0.13	9
Kleinfeld, Andrew J.	43	-0.14	9
Tatel, David S.	44	-0.22	DC
Seymour, Stephanie	45	-0.33	10
Rovner, Ilana	46	-0.33	7
Randolph, Arthur	47	-0.34	DC
Rogers, Judith	48	-0.36	DC
Karen J. Williams	49	-0.45	4
Frank M. Hull	50	-0.51	11
Edwards, Harry	51	-0.59	DC
Briscoe, Mary Beck	52	-0.62	10
Manion, Daniel	53	-0.63	7
Cabranes, Jose	54	-0.64	2
Kozinski, Alex	55	-0.68	9
Garland, Merrick B.	56	-0.72	DC
Sentelle, David Bryan	57	-0.74	DC
R. Guy Cole, Jr.	58	-0.84	6
James B. Loken	59	-0.85	8
Susan Harrell Black	60	-0.85	11
Joel Fredrick Dubina	61	-0.89	11
Parker, Fred I.	62	-0.95	2
Kanne, Michael	63	-0.99	7
Tacha, Deanell Reece	64	-1.03	10
Lucero, Carlos F.	65	-1.17	10
M. Blane Michael	66	-1.19	4

Table H Continued

Composite Ranking of Judges With Equal Weighting of Quality, Productivity, and Independence

(Includes only Active Circuit Court Judges 65 or Less in 2003)

Judges (Ranked highest to lowest scoring) based on equal weighting of Quality, Productivity, and Independence	(A) Rank based on Equal Weighted Composite Measure (Best = 1)	(B) Z-Score based on Equal Weighted Composite Measure	(C) Circuit
Barksdale, Rhesa	67	-1.19	5
Henry, Robert H.	68	-1.23	10
Henderson, Karen	69	-1.25	DC
Alice M. Batchelder	70	-1.25	6
Hawkins, Michael Daly	71	-1.34	9
Martha Daughtrey	72	-1.38	6
Rymer, Pamela Ann	73	-1.40	9
Stanley Marcus	74	-2.44**	11

** 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 20 Citation-Receiving Opinions)

Panel A: Seniority of Judges with Highest Top 20 Citation Counts

Judges with Highest Top 20 Outside Circuit Citation Count	Seniority Quartile on the Circuit	Chief Judge Status in 1998-2000	Years Experience
Sandra L. Lynch	4	0	3
Frank Easterbrook	2	0	13
Paul J. Kelly	2	0	6
Richard Posner	1	1	17
Bruce M. Selya	2	0	12
Anthony J. Scirica	1	0	11
Frank M. Hull	4	0	1
Karen J. Williams	3	0	6
Edward Earl Carnes	3	0	6
J. Harvie Wilkinson III	1	1	14
Mean	2.3	0.2	8.9

Panel B: Seniority of 11 Judges Centered on the Judge with the Median Number of Outside Circuit Citations for the Top 20 Opinions

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

Panel C: Seniority of Judges with Lowest Top 20 Citation Counts

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

Seniority Quartile 1 to 4 represent the top 25% to bottom 25% 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 ($p<0.221$)