

JUDGING THE TOURNAMENT

HON. JAY S. BYBEE* AND THOMAS J. MILES**

I. HOW ACCURATE ARE THE CHOI/GULATI MEASURES?.....	1058
A. <i>Productivity</i>	1058
B. <i>Quality</i>	1061
C. <i>Independence</i>	1065
D. <i>Other Dimensions</i>	1066
II. WHAT HAPPENS WHEN JUDGES “JUDGE TO THE TOURNAMENT?”.....	1068
III. SHOULDN’T WE COMPARE APPLES TO APPLES AND ORANGES TO ORANGES?.....	1073
IV. CONCLUSION.....	1076

The United States Constitution provides that the President has the power to appoint federal judges with the advice and consent of the Senate.¹ The Constitution does not specify the criteria that the President should use in selecting judicial nominees or that the Senate should employ in reviewing them. In recent years, the process of nominating and confirming candidates for the federal bench, and especially the Supreme Court, has become increasingly political and contentious. Professors Choi and Gulati criticize the apparently growing role ideology plays in choosing and evaluating judicial nominees and propose a bold alternative.² Their “Tournament of Judges” purportedly consists of a series of ideologically neutral measures that identify which appellate judges “merit” elevation to the Supreme Court.³ By restricting the choice of a nominee to the winner of the tournament, Professors Choi and Gulati hope to eliminate the role of ideology and the attendant partisan battling from the selection of Supreme Court Justices. Moreover, they claim that their market-based system for judicial selection would improve the quality of nominees.⁴ The current federal appellate bench, which is itself a product of the very system that Professors Choi and Gulati lament, should perhaps be grateful for their providing the equivalent of an HR manual for boosting each judge’s odds of promotion.⁵ But we are

* Judge, U.S. Court of Appeals for the Ninth Circuit.

** Olin Fellow in Law and Economics, The University of Chicago Law School. Miles is formerly a law clerk to Judge Bybee. The opinions expressed are solely those of the authors.

1. U.S. CONST. art. II, § 2, cl. 2.

2. See Stephen Choi & Mitu Gulati, *A Tournament of Judges?*, 92 CAL. L. REV. 299, 299 (2004) [hereinafter Choi & Gulati, *Tournament*]; Stephen J. Choi & G. Mitu Gulati, *Choosing the Next Supreme Court Justice: An Empirical Ranking of Judge Performance*, 78 S. CAL. L. REV. 23 (2004) [hereinafter Choi & Gulati, *Empirical Ranking*].

3. See Choi & Gulati, *Tournament*, *supra* note 2, at 301-02.

4. See *id.*; Choi & Gulati, *Empirical Ranking*, *supra* note 2.

5. Whether any judge should want to seek appointment to the Court is beyond Choi and Gulati’s study and our comments. Judge Frank Easterbrook has recently commented that “any judge who claims not to fancy a position on that Court is a liar.” Howard Bashman, *20 Questions for Circuit Judge Frank H. Easterbrook of*

convinced that evaluating judicial performance is not as easy as they suppose and that relying exclusively on the Tournament to select a Supreme Court nominee would not advance the rule of law.

As a preliminary comment, we applaud Professors Choi and Gulati for the purposes for which they undertake their study. Improving the quality of public discourse on the composition of the judiciary and increasing public awareness of the judiciary's work and how it functions are laudable goals.⁶ Moreover, empirical measurements, particularly empirical comparisons of a nominee to her judicial peers, can play an important role in informing decisionmakers and the public on the relative merit of a nominee. We disagree, however, that empirical measurements should be the sole basis on which a nominee should be chosen.⁷ In the end, the real mettle of a potential nominee to the Court lies in her opinions and character. Reading opinions (much less discerning character), however, is time-consuming and, hence, costly. It is, moreover, an inexact science. For members of the public without legal training, comprehending the often complex legal analysis of a judicial opinion is prohibitive. The public has turned instead to a less costly means of evaluating judicial nominees—looking at a nominee's positions on what Professors Choi and Gulati call the “hot button issues,” such as abortion, gun rights, affirmative action, and capital punishment, among others.⁸ Professors Choi and Gulati decry the hot-button approach as unduly narrowing the range of legal issues discussed and reducing the confirmation process “to quibbling over [a nominee's] expected position on issues like affirmative action and abortion.”⁹ They propose supplanting, not merely supplementing, the current nomination and confirmation process with their system of rankings.¹⁰ Their rhetorically charged use of the word “tournament” implies that the highest-ranked judge has won or earned a position on the Court and that the President and Senate should be reduced to the ministerial roles of simply awarding the tournament winner her rightful place on the Court.

An evaluation of whether the current nomination and confirmation system should be replaced with Choi and Gulati's tournament scheme requires comparing the costs and benefits of each and determining which offers society the greatest benefits net of costs. Natu-

the U.S. Court of Appeals for the Seventh Circuit, HOW APPEALING (Aug. 2, 2004), at http://www.legalaffairs.org/howappealing/20q/2004_08_01_20q-appellateblog_archive.html. We cannot vouch for Judge Easterbrook's view, but for purposes of our discussion, we will assume that whether or not any particular judge aspires to the Supreme Court, a rational judge aspires to be qualified for the Court.

6. See Choi & Gulati, *Tournament*, *supra* note 2, at 302-04.

7. See Choi & Gulati, *Empirical Ranking*, *supra* note 2, at 35-36.

8. See *id.* at 24, 34, 39.

9. *Id.* at 33.

10. See *id.* at 35-36.

rally, as advocates of the tournament, Professors Choi and Gulati lament the shortcomings of the current hot-button system because it fails to distinguish arguments about a nominee's merit from arguments about the nominee's ideology.¹¹ A politician's advocacy of, or opposition to, a particular nominee for purely "political" reasons appears unseemly, and politicians consequently attempt to mask their ideological arguments as arguments on the merits.¹² Implicit in Professor Choi and Gulati's justification for their tournament is that politicians have become so successful at conflating ideology and merit that the public can no longer distinguish the two.¹³ The current senatorial stalemate on many appellate nominees, which many view as a dress rehearsal for a confrontation over a Supreme Court nominee, is presumably one consequence.

In this Essay, we describe several potential shortcomings of using the tournament to select judicial nominees. Before we turn attention to these potential costs, some benefits of the hot-button approach are worth noting. First, hot-button issues are politically charged because they matter to the public. Other issues about which the public should be concerned—but is not—certainly exist. The public's failure to recognize the import of these other issues is a shortcoming of political discourse generally and is not specific to judicial confirmations; arguments over economic, domestic, and foreign policies are full of nuance and yet are frequently reduced to sloganeering and sound bites. Second, these issues may be hot buttons partly because they *do* have some power to predict a nominee's positions on other, more obscure legal issues. Third, if a Supreme Court nominee has written an appellate opinion on the hot-button issue, the opinion should prove highly informative of the nominee's abilities. When a judge confronts a case raising a hot-button issue, she knows that the public (and in the event of a nomination to the Supreme Court, the White House and Senate) will scrutinize the opinion, and she faces a strong incentive to put forth her best possible efforts. An appellate opinion on a hot-button issue should therefore represent the author's best judicial efforts and provide significant insights into her abilities as a legal analyst and stylist. In effect, each hot-button issue may constitute its own tournament of judges, the equivalent of a law review writing competition for judges.

The emphasis of this Essay is not that the current system lacks room for improvement—few would argue that it does not need improvement—but that the brave, new world of Supreme Court nominations proposed by Professors Choi and Gulati may not be as clearly

11. *Id.* at 34-36.

12. Choi & Gulati, *Tournament*, *supra* note 2, at 301.

13. *See* Choi & Gulati, *Empirical Ranking*, *supra* note 2, at 27-28, 38.

superior to the current system as they suppose. In this Essay, we discuss three concerns with using the Tournament of Judges as a basis for selecting a nominee to the Supreme Court. First, like other contributors to this Symposium, we question whether the metrics proposed by Professors Choi and Gulati appropriately measure the performance of circuit judges. Second, even if the Choi/Gulati metrics accurately capture judicial performance, the tournament itself may create incentives that distort judicial behavior and erode the quality of appellate judging. The criteria and the method by which a judge may improve her standing are readily known, and thus we are concerned that reducing judging to finite, measurable results would encourage judges to promote tournament criteria rather than adjudicate individual cases: judges may “judge to the test.” Third, even if the Choi/Gulati metric accurately measures the performance of circuit judges, winning the tournament may not predict success as a Supreme Court Justice.

I. HOW ACCURATE ARE THE CHOI/GULATI MEASURES?

A. *Productivity*

Professors Choi and Gulati measure judicial performance along three dimensions: (1) productivity, (2) quality, and (3) independence.¹⁴ For productivity, they tally the number of opinions each judge published, because they assume that unpublished opinions “often involve minimal effort (and a lower quality of reasoning).”¹⁵ However, norms about when a case warrants an opinion differ across circuits. Professors Choi and Gulati correct for these differing norms by adding to each judge’s tally the difference between the average number of opinions published in the judge’s circuit and the average published in the most productive circuit, which, currently, is the Seventh Circuit.¹⁶ The appeal of this measure is that a judge who resolves more cases in published opinions appears hardworking, and it may often be the case. However, the measure is not without difficulties.

First, Professors Choi and Gulati’s assumption that unpublished opinions involve a lower quality of reasoning is dubious. In the Ninth Circuit, the test for whether to publish a decision is whether existing precedent squarely controls the issues presented.¹⁷ A judge’s failure

14. *Id.* at 42.

15. *Id.* at 43.

16. *Id.* at 45.

17. See 9TH CIR. GEN. ORDERS 4.3.a (“A memorandum disposition cannot be cited as precedent. Unlike an opinion for publication which is designed to clarify the law of the circuit, a memorandum disposition is designed only to provide the parties and the district court with a concise explanation of this court’s decision. Because the parties and the district court are aware of the facts, procedural events and applicable law underlying the dispute, the disposition need recite only such information crucial to the result.”); see also 9TH

to see novel issues in a case may indicate deficiencies in the judge's intellectual curiosity and legal acumen, or it may reflect the judge's respect for existing precedent. Cases vary in the novelty of the legal questions they present and, hence, in the degree of guidance the parties and future litigants require. As the Ninth Circuit rules indicate, cases that lack novelty should not be published because they offer current and future parties little or no instruction beyond existing precedent. In addition, judicial resources are constrained, at the very least, by the judge's time. The opportunity cost of the time that a judge spends preparing a full opinion on a case furnishing little instruction to the public is the time that she could spend on another case presenting unresolved questions of law. Professors Choi and Gulati implicitly assume that all cases offer the same amount of instruction and the same amount of social benefit because they believe that all cases deserve published opinions. In their tournament, this assumption becomes an incentive to maximize publication. If pursued, this objective would work a substantial reallocation of judicial labor to less pressing questions.¹⁸ The tournament encourages a substitution that would decrease rather than increase the amount of social benefit that the appellate courts provide.

Second, Choi and Gulati's focus on published opinions is not devoid of ideological content. It disfavors advocates of judicial restraint who may prefer fewer published pronouncements from the bench. A recent study argues that the ideological leaning of a three-judge panel correlates with the decision to publish.¹⁹ The very purpose of

CIR. R. 36-2 ("A written, reasoned disposition shall be designated as an OPINION only if it: (a) Establishes, alters, modifies or clarifies a rule of law, or (b) Calls attention to a rule of law which appears to have been generally overlooked, or (c) Criticizes existing law, or (d) Involves a legal or factual issue of unique interest or substantial public importance, or (e) Is a disposition of a case in which there is a published opinion by a lower court or administrative agency, unless the panel determines that publication is unnecessary for clarifying the panel's disposition of the case, or (f) Is a disposition of a case following a reversal or remand by the United States Supreme Court, or (g) Is accompanied by a separate concurring or dissenting expression, and the author of such separate expression requests publication of the disposition of the Court and the separate expression.")

18. See RICHARD A. POSNER, *THE FEDERAL COURTS: CHALLENGE AND REFORM* 168-69 (1996) ("Given the workload of the federal courts of appeals today, the realistic choice is not between limited publication, on the one hand, and, on the other, improving and then publishing all the opinions that are not published today; it is between preparing but not publishing opinions in many cases and preparing no opinions in those cases."); *Hart v. Massanari*, 266 F.3d 1155, 1177 (9th Cir. 2001) ("[F]ew, if any, appellate courts have the resources to write precedential opinions in every case that comes before them."); *id.* at 1179 ("Adding endlessly to the body of precedent . . . can lead to confusion and unnecessary conflict.")

19. See David S. Law, *Strategic Judicial Lawmaking: Ideology, Publication, and Asylum Law in the Ninth Circuit*, 73 U. CIN. L. REV. 817 (2005) (arguing that an examination of only published opinions leads to the erroneous conclusion that panels dominated by judges appointed by Republican Presidents were as likely to rule in favor of asylum seekers as panels dominated by judges appointed by Democratic Presidents).

the tournament is to develop a measure of judicial merit free of ideology, but by examining only published opinions, the tournament retains ideological content. It favors judges who distinguish each case from existing precedent and thereby foster complexity in the law.

Third, cases vary widely in their difficulty, but not in the Choi/Gulati rankings. In the tournament, a judge who resolves many small cases presenting simple issues in short opinions will rank higher than a judge who resolves a few large cases presenting complex issues in long opinions. In reality, the latter judge may have exerted greater effort and possess superior judicial talents. Fourth, even if the Choi/Gulati metric reflected pure productivity, intra- and intercircuit comparisons are a troublesome complication. In this measure, a judge who publishes more opinions than the average judge on his court and who sits in a circuit where the average number of opinions published per judge is above that of other circuits will be ranked higher than a judge who publishes fewer opinions than the average judge on his court and who sits on a circuit where the average number of opinions published per judge is below that of other circuits. However, the middling cases are less clear-cut. Is a judge who publishes more opinions than the average judge on his court but who sits in a circuit where the average number of opinions published per judge was below the average of other circuits better or worse than a judge who publishes fewer opinions than the average judge on his court but who sits in a circuit where the average number of opinions published per judge was above the average of other circuits? It is not obvious whether an above-average judge on a below-average circuit is better than a below-average judge on an above-average circuit. Without justification, Professors Choi and Gulati favor the former. They state that “[a]ny differences among judges will, therefore, be determined solely by each judge’s standing relative to the other judges within her own circuit,”²⁰ and their productivity measure thus favors judges who are stars on their circuits. If their measure applied to basketball, Scottie Pippin playing for the Portland Trail Blazers would fare better than Scottie Pippin playing for the Chicago Bulls. In their metric, the presence of other all-stars who raise the circuit’s average productivity make it harder for a star judge to shine. Notably, judges on the D.C. Circuit, widely considered an all-star circuit and the warm-up bench for the Supreme Court, fare poorly in the Choi/Gulati productivity metric.²¹

20. Choi & Gulati, *Empirical Ranking*, *supra* note 2, at 46.

21. *See id.* at 76-77.

B. Quality

Given their view of the current system of confirmation, Professors Choi and Gulati use a measure of quality that is inconsistent with their assumptions about judicial behavior. The tournament assumes that the number of citations an opinion receives from courts outside its circuit reflects its quality.²² Yet Professors Choi and Gulati believe that many nominees are advanced on the basis of ideology rather than merit.²³ The tournament does not consider the possibility that the citations may themselves be objects of ideological manipulation rather than indications of the quality of the legal analysis.²⁴ For citations to reflect solely quality, judges would have to shed their ideological leanings once they reach the bench. If judges did so, an ideological President would have no reliably ideological judges to nominate to the Court. Were it so, the problem of ideologically driven nominations would vanish, and the tournament itself would be unnecessary.

Even if judges do not choose their citations on the basis of ideology, Professors Choi and Gulati's measure of quality may actually reflect other characteristics. They acknowledge that judges differ in the initial probability that their opinions are read.²⁵ The notoriety a judge enjoyed before joining the bench, or enjoyed while on the bench but before the observation period of this study, carries over onto the study's observation window. A lawyer known to legal commentators is more likely to have her opinions read once she takes to the bench, and hence she has a greater probability of having her opinions cited than does a judge whose prior legal career was relatively obscure. Professors Choi and Gulati claim that "it is precisely this [pre-tournament] reputation for quality analysis (and the underlying ability behind such a reputation) that we hope to find in a Supreme Court nominee."²⁶ The criterion favors academics over practitioners because publishing more readily establishes reputations than does winning cases or structuring deals. This measure promotes, for example, a Scalia, Ginsburg, or Breyer over a Powell, Blackmun, or Souter. Moreover, the labor force of the judiciary gives erstwhile (or part-time) legal academics on the bench a further advantage. Law clerks who draft opinions are typically recent graduates of law schools, where they studied the casebooks and commentaries of these

22. See *id.* at 48-50.

23. See *id.* at 30-31.

24. A recent study purports to find strong evidence of ideological voting in three-judge appellate panels. See Cass R. Sunstein et al., *Ideological Voting on Federal Courts of Appeals: A Preliminary Investigation*, 90 VA. L. REV. 301 (2004).

25. See Choi & Gulati, *Empirical Ranking*, *supra* note 2, at 49. "Our methodology does not control for such pretournament inherent differences." *Id.*

26. *Id.*

judges qua academics. Having previously relied on the instruction of these professors' writings, law clerks are likely to look to the opinions of these familiar oracles for guidance. Furthermore, a judge's opinions may receive attention for reasons unrelated to his or her reputation for quality. The events that generate interest in a judge's opinions may or may not stem from the trenchancy of his legal reasoning on the bench.²⁷

The tournament's quality measure may also reflect speed in rendering a decision on a common question faced by all circuits. Witness the recent flurry of opinions following the Supreme Court's decision in *Blakely v. Washington*.²⁸ The *Blakely* court held that the criminal sentencing guidelines of Washington State violated the Sixth Amendment because they did not require that every fact raising a sentence beyond the statutory maximum be proven to a jury beyond a reasonable doubt or admitted to by the defendant.²⁹ Immediately, commentators and litigants wondered if the logic of *Blakely* would apply to the federal sentencing guidelines. Twelve days after the Supreme Court issued *Blakely*, the Seventh Circuit heard oral argument in *United States v. Booker*,³⁰ and three days later, the court issued a decision. The majority in *Booker* held that *Blakely* implied that sentence enhancements under the U.S. Sentencing Guidelines based on facts not determined by a jury beyond a reasonable doubt violated the Sixth Amendment.³¹ Twelve days after *Booker*, the Ninth Circuit decided *United States v. Ameline*,³² in which the majority largely agreed with the Seventh Circuit: "We join the Seventh Circuit in holding that there is no principled distinction between the Washington Sentencing Reform Act at issue in *Blakely* and the United States Sentencing Guidelines."³³

In the six weeks following the respective decisions, courts in other circuits cited *Booker* twenty-eight times but cited *Ameline* only thirteen times. Despite a time lag of less than two weeks, the earlier decision was cited twice as many times. The advantage in terms of subsequent citations to the first court to resolve an issue is significant. Each circuit court looks to its sister circuits in confronting legal issues, in part because another circuit's analysis may offer guidance.

27. See, e.g., Article III Groupie, *Judicial Divas and Hotties*, UNDERNEATH THEIR ROBES: NEWS, GOSSIP, AND COLORFUL COMMENTARY ABOUT THE FEDERAL JUDICIARY (Nov. 1, 2004), at http://underneaththeirrobes.blogspot.com/main/judicial_divas_and_hotties/index.html (treating Article III judges as celebrities).

28. 124 S. Ct. 2531 (2004).

29. See *id.* at 2537-38.

30. 375 F.3d 508 (7th Cir. 2004), *aff'd and remanded by* 125 S. Ct. 738 (2005).

31. See *id.* at 510, 513, 515.

32. 376 F.3d 967 (9th Cir. 2004), *amended and superseded on reh'g*, 400 F.3d 646 (9th Cir. 2005), *reh'g en banc granted*, 401 F.3d 1007 (9th Cir. 2005).

33. *Id.* at 974 (footnote omitted).

In addition, a decision that conflicts, rather than conforms, with that of another circuit is more likely to be called en banc or to prompt a certiorari grant. The first circuit to weigh in on a legal issue is therefore more likely to be cited than subsequent circuits that agree with its conclusion. The tournament's quality measure may reflect this first-mover advantage rather than which opinion contains the most thorough analysis.

The first answer is not necessarily the right one. Substantively, the circuits split on *Blakely's* relevance to the federal guidelines, and after granting certiorari in *United States v. Booker*, the Supreme Court recently resolved the split.³⁴ The legal community viewed several aspects of the Court's holding as a "surprise," namely, the conclusions that judges must consult with the guidelines but that they are only advisory, and that sentences are subject to reasonableness review on appeal. Neither the author of the Seventh Circuit *Booker* majority (Judge Posner) nor its dissenter (Judge Easterbrook) nor any other appellate judge chose the same remedy as the Supreme Court did. The failure of any circuit judge to anticipate the Supreme Court's holding during the few months between the *Blakely* decision and the grant of certiorari in *Booker* illustrates that fast analysis is not always the same as the Supreme Court's analysis.

Another consideration is that opinions with clever turns of phrase are cited more often. A well-known rule of appellate practice is that an appellant's brief must contain the argument with the "appellant's contentions" and the reasons for them, with citations to the authorities and parts of the record on which the appellant relies.³⁵ One purpose of this rule is to economize on judicial resources by sparing judges the task of searching for litigant's arguments. In *United States v. Dunkel*,³⁶ the Seventh Circuit vividly expressed this rule: "Judges are not like pigs, hunting for truffles buried in briefs."³⁷ We found thirteen instances in which other circuit courts quoted this line in published opinions.³⁸ We located another two unpublished opinions

34. 125 S. Ct. 738 (2005).

35. FED. R. APP. P. 28(a)(9)(A).

36. 927 F.2d 955 (7th Cir. 1991) (per curiam).

37. *Id.* at 956.

38. See *Crossley v. Georgia-Pacific Corp.*, 355 F.3d 1112, 1114 (8th Cir. 2004); *Malacara v. Garber*, 353 F.3d 393, 405 (5th Cir. 2003); *Coggin v. Longview Indep. Sch. Dist.*, 337 F.3d 459, 468-69 (5th Cir. 2003) (en banc) (Jones, J., dissenting) (noting "[t]he problem has been colorfully, if hyperbolically, described by our brethren on the Seventh Circuit"); *Indep. Towers of Wash. v. Washington*, 350 F.3d 925, 929 (9th Cir. 2003); *Roska ex rel. Roska v. Peterson*, 328 F.3d 1230, 1246 n.13 (10th Cir. 2003); *Craven v. Univ. of Colo. Hosp. Auth.*, 260 F.3d 1218, 1226 (10th Cir. 2001); *United States v. Stuckey*, 255 F.3d 528, 531 (8th Cir. 2001); *Downs v. L.A. Unified Sch. Dist.*, 228 F.3d 1003, 1007 n.1 (9th Cir. 2000) (paraphrasing); *Gross v. Burggraf Constr. Co.*, 53 F.3d 1531, 1546 (10th Cir. 1995); *Teague v. Bakker*, 35 F.3d 978, 985 n.5 (4th Cir. 1994); *Enplanar, Inc. v. Marsh*, 11 F.3d

in other circuits that cite *Dunkel* for the proposition that cursory allegations are inadequate but do not quote the pigs-and-truffles language directly.³⁹ An earlier case in the Seventh Circuit made the same point: Rule 28 implies that appellate courts are not responsible for scouring briefs and accompanying documents for poorly articulated and supported arguments. However, it expressed the point less colorfully, and other circuit courts have not cited it for this proposition. In *United States v. Williams*, a criminal defendant appealed his conviction in part on the ground that the trial court improperly excluded hearsay testimony, but the appellant failed to identify the portions of testimony he sought to have admitted at trial.⁴⁰ The court held that the appellant's conclusory allegation did not comply with Rule 28 and was therefore waived.⁴¹ It also noted, as it did in *Dunkel* but in less memorable language, that the court (and the prosecutor) are not responsible for locating the portions of the record relevant to the appellant's claim: "Neither this court nor the United States Attorney has a duty to comb the record in order to discover possible errors."⁴² We have found no published opinion in another circuit that cites this language or even cites *Williams* for this proposition. Clever turns of phrase can be helpful mnemonics or just plain amusing, but they do not necessarily indicate deeper legal reasoning. In fact, most legal writing texts instruct that workmanlike prose and the avoidance of novel turns of phrase best achieve the clarity valued in legal analysis.⁴³ Although writing ability and legal reasoning are surely complementary skills, judges' appetite for the one-liners of their colleagues implies that a few zingers may distort the tournament's measure of "quality."

1284, 1297 n.16 (5th Cir. 1994); *Greenwood v. FAA*, 28 F.3d 971, 977 (9th Cir. 1994); *Sanjour v. EPA*, 984 F.2d 434, 437 n.3 (D.C. Cir. 1993).

39. See *El-Fadl v. Cent. Bank of Jordan*, 75 F.3d 668, 675 (D.C. Cir. 1996); *Larson v. Nutt*, 34 F.3d 647, 648 (8th Cir. 1994) (per curiam).

40. 877 F.2d 516, 518 (7th Cir. 1989).

41. *Id.* at 518-19.

42. *Id.* at 519.

43. See BRYAN A. GARNER, *THE ELEMENTS OF LEGAL STYLE* 6-8 (2d ed. 2002) (emphasizing clarity); HENRY WEIHOFEN, *LEGAL WRITING STYLE* 61-82 (2d ed. 1980) (discussing the goal of simplicity). Although it is not generally recognized as a legal writing text, Dickens' *David Copperfield* is instructive on the matter:

'How do you like the law, Mr. Micawber?'

'My dear Copperfield,' he replied. 'To a man possessed of the higher imaginative powers, the objection to legal studies is the amount of detail which they involve.

Even in our professional correspondence,' said Mr. Micawber, glancing at some letters he was writing, 'the mind is not at liberty to soar to any exalted form of expression. Still, it is a great pursuit. A great pursuit!'

CHARLES DICKENS, *DAVID COPPERFIELD* 524-25 (Modern Library Paperback ed. 2000) (1850).

C. Independence

Professors Choi and Gulati also attempt to measure whether a judge thinks independently rather than ideologically. They measure independence as the frequency with which a judge opposes a fellow panel member appointed by a President of the same political party, relative to the frequency with which the judge sits with judges of the same political party.⁴⁴ The assumption of this measure is that an ideological judge is less likely to dissent from decisions joined by members of her own party, and an independent judge is as likely to dissent from a decision joined by members of her own party as she is to dissent from one joined by members of the other party. This model of judicial decisionmaking, that a judge's political tendencies mirror those of her appointing President, is naturally an oversimplification. However, even if this assumption were generally valid, Professors Choi and Gulati's measure misses an important dimension of independence. Under their assumption, a judge's independence from her appointing President could be measured in two ways: the frequency with which a judge dissents from members of her own party and the frequency with which she agrees with members of the opposing party. However, the tournament considers only the former and thereby risks mismeasurement. For example, a highly ideological judge who outflanked members of her own party on the ideological extreme might regularly dissent from decisions joined by members of her own party because she perceived them as too moderate. In so doing, the highly ideological judge would appear highly independent in the tournament. By not considering the regularity with which the judge dissented from members of the opposing party, the tournament may fail to identify a judge who is persistently political.

This hypothetical might be criticized as too far-fetched, but it also points out the conceptual inconsistency in the tournament's quality and independence measures. For the quality measure to reflect the superiority of legal reasoning and analysis, judges must cite opinions on the basis of their objective worth, independent of their politics. For the tournament's independence measure to reflect autonomous thinking, judges must be predictably political in deciding cases. The models of judicial behavior underlying the two measures are in tension with one another.

The tournament's independence measure fails to account for other complexities of judicial behavior. A judge who thinks independently might be more willing to compromise and may dissent relatively infrequently. For that judge, a small change in the number of her dissents would greatly influence how she performed along this measure.

44. See Choi & Gulati, *Empirical Ranking*, *supra* note 2, at 61-67.

Similarly, judges who stave off writing dissents by persuading their peers of their own view may be highly independent thinkers, but they would not appear so in this measure. In addition, a judge sitting in a circuit that consistently renders the legally “correct” decision could excel in this measure only by producing dissents of questionable value.

D. Other Dimensions

The tournament is hampered by its exclusion of certain characteristics conventionally thought to be important in judges and Justices. A judge may contribute mightily to the quality of an opinion even if she is not its author. A thoughtful judge may ask penetrating questions from the bench that help shape the views of the other members of the panel. In conference discussions or in commenting on a colleague’s draft opinion, a judge may influence an opinion’s analysis. In circuits with frequent en banc rehearings, a judge may spend significant energy reviewing the court’s decisions and writing to inform her colleagues of the desirability of rehearing particular cases en banc, but the public is never informed of these efforts. These characteristics and efforts are reflected only obliquely in the tournament to the extent that they correlate with the three dimensions measured.

To a good degree, however, all of these criticisms are snipping around the edges of the Choi/Gulati tournament. What is remarkable about their exercise is the dominance of Judges Posner and Easterbrook along the productivity and quality measures, and the bunching together of virtually everyone else. Although the Seventh Circuit duo measures up less well in the independence metric, persons who are at all familiar with their writings—from on or off the bench—would not question the independence of their thinking. Their performance in this category may itself indicate that the quality of independence is extremely difficult to quantify. Their dominance also raises the question of why the tournament is necessary at all. Everyone in the legal community knows that these two judges are the rarest of talents. The question that Choi and Gulati do not ask is why these two do not belong to the so-called “Bush Five,” the rumored short-list of potential nominees. The fact that they are not on the rumored list suggests that the tournament fails to capture some relevant phenomenon.

In their writings off the bench, Judges Posner and Easterbrook have at times been intellectual provocateurs. For example, Judge Easterbrook has criticized mandatory disclosure rules in the federal securities laws⁴⁵ and has characterized—not disapprovingly—the

45. See Frank H. Easterbrook & Daniel R. Fischel, *Mandatory Disclosure and the Protection of Investors*, 70 VA. L. REV. 669, 714 (1984) (“We have not constructed a compelling

rules of criminal procedure as creating a marketplace.⁴⁶ Judge Posner has explored the possible benefits of baby-selling⁴⁷ and licenses to commit rape,⁴⁸ to name just two examples. These academic writings are noteworthy, even classic, pieces of legal scholarship because they rigorously analyze why the improbable could be possible. A penchant for provocation, sometimes at the expense of an audience's sensibilities, makes for interesting reading, classroom discussion, and academic symposia, but it might not be a desirable characteristic in a Supreme Court Justice.

The public might prefer its judges and Justices to possess numerous other characteristics that the tournament excludes. For example, the public might want its judges to be fair, genuinely concerned about the parties coming before the court, and humbled by the authority entrusted to them.⁴⁹ These qualities—often lumped together under the rubric “judicial temperament”—do not readily lend themselves to quantification. The tournament's best performers, as well as its worst ones, might have these characteristics in abundance, or they may have them in varying quantities. Because the tournament does not include them, the reader simply does not know.

The bunching together of other judges and the observation that small changes in the number of opinions, citations, or dissents significantly reorder the rankings suggest that the tournament lacks something else. Specifically, it does not contain a measure that permits most judges to be distinguished from one another.⁵⁰ This absence may be an inadequacy of the tournament or it may reflect the technology of judging. Perhaps most appellate judges perform at roughly the same level because they face the same technology, possess the same staff resources, and, within a certain range, are armed with the same amount of intellectual firepower. Perhaps the most important similarity in the work of federal judges is that all of them labor with and under the same body of law: the Constitution, federal statutes and regulations, and precedent of the Supreme Court and the respective circuit courts.

The appropriate interpretation of the tournament's main result—that, with few exceptions, most federal judges perform roughly

case for [securities] regulation of any sort, let alone for the particular regulations the SEC uses.”).

46. See Frank H. Easterbrook, *Criminal Procedure as a Market System*, 12 J. LEGAL STUD. 289 (1983).

47. See Elisabeth M. Landes & Richard A. Posner, *The Economics of the Baby Shortage*, 7 J. LEGAL STUD. 323 (1978).

48. RICHARD A. POSNER, *ECONOMIC ANALYSIS OF LAW* § 7.1, at 216-17 (6th ed. 2003).

49. See, e.g., Brett Scharffs, *The Role of Humility in Exercising Practical Wisdom*, 32 U.C. DAVIS L. REV. 127 (1998).

50. See Choi & Gulati, *Empirical Ranking*, *supra* note 2, at 75 (“Without Judge Posner or Judge Easterbrook, no one winner emerges.”).

equally—is unclear. Professors Choi and Gulati chose to emphasize a few judges at the top of the ranking rather than the ranking’s inability to distinguish most of the judges. The primary lesson of the tournament might instead be that since one federal judge is about as good as another along the dimensions they measure, the few outliers in the tournament are effectively statistical noise. If so, this interpretation would argue in favor of replacing the current nomination process with a randomization device in which a Supreme Court nominee was chosen by lot from the pool of federal appellate judges. A randomization process would achieve their goal of reducing partisan bickering. Curiously, Professors Choi and Gulati did not consider this alternative.

Ultimately, the exclusion from the tournament of relevant characteristics imparts some measurement error to the prediction of the judge most qualified to sit on the Supreme Court. This inaccuracy implies that a nomination of a judge other than the tournament winner to the Supreme Court may not, as Professors Choi and Gulati contend, be solely a function of ideology.

II. WHAT HAPPENS WHEN JUDGES “JUDGE TO THE TOURNAMENT?”

Professors Choi and Gulati emphasize that the tournament measures a judge’s productivity or work effort.⁵¹ By calling their rankings a tournament, they reference an economics literature in which hierarchical reward structures induce greater effort from workers.⁵² The suggestion that appellate judges need competition implies that the current system is not operating optimally. But the idea that appellate judges exert insufficient amounts of effort contrasts with the prevailing wisdom that the federal appellate bench is understaffed. Since 1960, the number of cases filed in federal appellate courts rose by more than fifteenfold, but the number of judges increased by only 2.5 times.⁵³ As a result, the average active judge handles 956 more cases per year than she did forty years ago.⁵⁴ Although caseload var-

51. See Choi & Gulati, *Empirical Ranking*, *supra* note 2, at 43 (describing published opinions as requiring more effort than unpublished ones); *id.* at 62 (emphasizing the effort required to write dissents and concurrences).

52. See Sherwin Rosen, *Prizes and Incentives in Elimination Tournaments*, 76 AM. ECON. REV. 701, 709 (1986) (explaining that prizes in a sequential elimination tournament must be concentrated at the top in order to induce sufficient effort in later rounds); *cf.* Edward P. Lazear & Sherwin Rosen, *Rank-Order Tournaments as Optimum Labor Contracts*, 89 J. POL. ECON. 841, 849 (1981) (suggesting that a tournament induces optimal investment in skills).

53. ADMIN. OFFICE OF THE U.S. COURTS, JUDICIAL FACTS AND FIGURES: MULTI-YEAR STATISTICAL COMPILATIONS OF FEDERAL COURT CASELOAD THROUGH FISCAL YEAR 2004, at tbl.1.3 (2005), available at <http://www.uscourts.gov/judicialfactsfigures/table1.03.pdf>.

54. See *id.*

ies considerably across circuits, it has risen in all circuits. The crush of cases already provides judges ample incentive to work hard.

Competition among circuit judges is a curious virtue to extol. If the overall effort of appellate judges is not in question, then the need for competition suggests that the distribution of judicial resources has been misallocated and that the tournament would redirect judicial resources in the right direction. It is here that we have our greatest differences with Choi and Gulati's proposal. It is precisely because the tournament's qualifying criteria are known to each judge that there is a possibility of competition. We may fairly ask: Do we really wish to foster competition along the productivity, quality, and independence measures? Are there perverse consequences to the competition? What happens when judges "judge to the tournament"?

The structure of the federal appellate courts appears designed to discourage competition among judges. In contrast to the separated branches, which face "hydraulic pressure . . . to exceed the outer limits of [their] power,"⁵⁵ *within* the judicial branch appellate courts are collegial, not competitive, bodies. Circuit courts have a near monopoly over federal cases and controversies within their jurisdiction, limited only by the Supreme Court's decisions on those questions. Collegial decisionmaking not only encourages but demands collusion (in the economic sense) among judges. The terms of employment for federal judges discourage their thinking about possible promotion to the Supreme Court. Federal appellate judges are not even guaranteed consideration for a nomination because the President may look beyond the federal bench for good candidates. Perhaps most important, Article III of the Constitution grants each judge a lifetime appointment, conditional on good behavior.⁵⁶ Life tenure spares a judge the temptation to tailor her decisions to retain her current position or to angle for the next one. Professors Choi and Gulati would undo this system and have judges altering their decisions to jockey for an appointment to the Court.

The first "competitive" consequence of the tournament would be analogous to the "school-qualifying-exam" effect. Exam-based promotion systems in secondary schools have as their goal that students demonstrate minimal competency. Critics contend, however, that the tests create an incentive for educators to maximize passage rates by

55. *INS v. Chadha*, 462 U.S. 919, 951 (1983).

56. U.S. CONST. art. III, § 1. Life tenure in a single position contrasts with the usual economic rationale for a tournament. Often an actor's skills are unknown, and a sequential-elimination tournament provides a cost-effective mechanism for players to determine if they possess the skills necessary to succeed in the endeavor. If they do not, they may quickly exit and seek other pursuits. See Lazear & Rosen, *supra* note 52, at 861 (suggesting "tryouts" when ability is unknown); Glenn M. MacDonald, *The Economics of Rising Stars*, 78 AM. ECON. REV. 155 (1988).

ignoring any subject not examined. Similarly, ambitious judges might ignore any activities and abandon any qualities that do not advance their rankings in at least one of the three dimensions of the tournament. For the reasons previously described, this single-mindedness would not serve the public well. Second, even if judges limited their attention to just the criteria of the tournament, they might still modify their decisionmaking in ways that do not further the rule of law or the quality of justice. Perhaps the easiest dimension to manipulate is the productivity ranking. A judge's relative position improves by designating more decisions for publication. A judge need not actually improve the quality of the decision in order to advance in the tournament's productivity ranking. Unlike law professors who must persuade journal editors that their writings are worth publishing, *Federal Reporter, Third Series*, is a circuit judge's Frostian home: when judges go there, West Publishing has to take them in. A counterargument may be that a judge's concern with the quality of her opinions, as measured either by the tournament or by her reputation in the legal community, constrains her willingness to publish everything. This argument has some merit, but it falls short in three ways.

First, a recent proposed change in the Federal Rules of Appellate Procedure (F.R.A.P.) has prompted circuits with relatively large unpublished opinion practices to consider how they would respond if every decision were treated as a published decision, that is, as precedent that can be cited to the court.⁵⁷ The conventional wisdom is that if proposed F.R.A.P. 32.1 were adopted, then a decision, which previously would have contained some analysis and explanation of the decision for the benefit of the parties, will instead issue as a summary affirmance or reversal, albeit a published one. The tournament redoubles the incentive for summary decisions. Because the tournament counts only published opinions without reference to their content, summary decisions boost a judge's productivity ranking even though they require less effort than the types of unpublished decisions issued currently.

Second, the quality rankings interact with the incentive to publish in odd ways. Citation counts encourage judges to distinguish prior precedents, even to author conflicting opinions, since such opinions will likely be acknowledged by other courts. Take our prior example of the Seventh Circuit's *Booker* decision. A court is more likely to cite *Booker* itself than a subsequent opinion that agrees with *Booker* and

57. One of us has publicly commented on the proposed change, opposing it. Letter from Jay S. Bybee, Judge, U.S. Court of Appeals for the Ninth Circuit, to Peter G. McCabe, Secretary, Committee on Rules of Practice and Procedure (Feb. 10, 2004), available at http://www.secretjustice.org/pdf_files/Comments/03-AP-327.pdf.

adopts its approach. However, an opinion that disagrees with *Booker* or that adds a nuance to its analysis is more worthy of citation. An ambitious judge in another circuit would therefore have an incentive to depart in some way from *Booker*, even if she agreed with its reasoning and conclusion. By relying on citation counts, the tournament rewards judges for the novelty of their opinions; the result may be less clarity or certainty in the law.

Third, one version of the quality rankings counts just the citations to a judge's "top twenty" opinions.⁵⁸ This scheme strengthens the incentive to issue summary judgments in cases that do not present novel or interesting questions. An ambitious judge could summarily decide the bulk of cases and focus her attention on crafting a few spectacular opinions. In so doing, she would boost her tally of published opinions and jumpstart her quality ranking. The quality rankings create incentives for other strategic behaviors. As described above, the tournament's quality measure spurs pithy statements and judicial aphorisms. This reward structure is at some tension with the actual quality of legal analysis. In the political sphere, one-liner characterizations of a candidate's positions are much decried. The tournament rewards the proliferation of this practice in judicial writing, where more thoughtful analysis should prevail. Furthermore, a self-interested judge might collude with other judges to boost her citations. Senior judges who visit other circuits have an opportunity to cite their home-circuit peers in the opinions they author in the visited circuit. The tournament creates an incentive for a judge seeking to boost her quality score to engage in "trades" with a senior judge in exchange for out-of-circuit citations.

Even if a judge cannot increase her own quality rankings, she can suppress those of her rivals in other circuits. The transparency of the criteria imply that a judge knows precisely the identity of the other judges clustered around her and, thus, whose opinions to avoid citing. The tournament creates an incentive for judges to do the opposite of what Professors Choi and Gulati believe makes their quality measure plausible: to cite opinions for strategic reasons rather than for the quality of the legal reasoning and analysis. The tournament itself corrodes the validity of this measure.

The tournament's independence ranking may have the most predictable incentive effect. To stand out in this measure, a judge should simply dissent more often when she sits with judges of the same political party. Additional dissents and concurrences breed confusion and complexity in the law. The American courts have a tradition of fostering collegial opinions. Among Chief Justice Marshall's greatest contributions to the Supreme Court was his leadership in effecting

58. See Choi & Gulati, *Empirical Ranking*, *supra* note 2, at 54-58.

its departure from the English practice of *seriatim* opinion writing.⁵⁹ The tournament erodes this tradition. A greater agreement among the panel indicates the judiciary's solidarity in the basis of its judgment. We recognize this when we debate the viability of 5-4 decisions of the Court, and we acknowledge the unassailability of a unanimous judgment and a single opinion. Perhaps the strongest signal of solidarity that a court can communicate is the unanimous and unsigned *per curiam* opinion. The tournament weakens the willingness of judges to join a *per curiam* because doing so forgoes an opportunity to appear independent. It also weakens the incentive to draft a *per curiam* because it denies the author an additional point in the productivity rankings.

Our criticism of the tournament should not be mistaken for mere cynicism. By suggesting that judges might engage in strategic behavior in response to the tournament, we do not impugn the integrity of the American bench. Our criticism is directed to the criteria for winning the tournament, which, Professors Choi and Gulati tell us, are conspicuously "transparent"⁶⁰ precisely so that judges will know what they have to do to win, and "gaming is a good thing."⁶¹ What we see as an objection—that judges will compete in the tournament—Professors Choi and Gulati see as the tournament's virtue. However, as we have described, the tournament creates incentives for a gamesmanship that does not reflect desirable qualities in a judge or Justice, and therefore a whole new set of additional rules might have to be introduced to ensure the tournament actually represents the "fair play" that its creators envision.

The establishment of the tournament as a selection device for Supreme Court nominees also changes the incentives facing the President. Professors Choi and Gulati intend the tournament to motivate the President to pick its winner.⁶² However, the greater the import of the tournament, the greater the temptation to pick a candidate from outside the federal appellate bench. Sidestepping the tournament altogether is the same maneuver that some Presidents have purportedly taken to eschew scrutiny of a nominee's views on hot-button issues. As previously described, a judicial opinion on a hot-button issue may provide insights into a nominee's ability. A candidate who has not written on these issues is more difficult to evaluate, and hence to oppose, because relatively less is known about her. Similarly, a candidate who is unranked in the tournament cannot be compared to the

59. See R. KENT NEWMYER, JOHN MARSHALL AND THE HEROIC AGE OF THE SUPREME COURT 157 (2001); G. EDWARD WHITE, THE MARSHALL COURT AND CULTURAL CHANGE 1815-1835, at 191 & n.145 (1991).

60. Choi & Gulati, *Empirical Ranking*, *supra* note 2, at 36.

61. *Id.* at 34.

62. See *id.* at 27-28.

set of alternative candidates. By picking a nominee who is not ranked in the tournament, a President could avoid precisely the scrutiny that the tournament is intended to apply. Rather than expanding the amount of information available about nominees, the tournament encourages the selection of “stealth” candidates or candidates without paper trails.

The tournament also alters the public’s incentives. The ranking of judges creates a new set of data easily grasped by the public. A President who nominates number forty-two on the list will have to explain himself.⁶³ And even if the top ten on the list are not feasible candidates for one reason or another, the President will still have to explain why judges ranked eleven through forty-one are not available for nomination. In contrast, he will not have to explain why he did not nominate the judge ranked forty-third because the reason will be obvious to everyone. That is a heavy responsibility to place on a ranking. Additionally, the content-free nature of the tournament makes it easier for the public to have an “informed” opinion about a nominee’s fitness for the Court without acquiring information about the nominee’s judicial philosophy. Like baseball fans who read box scores and participate in rotisserie leagues but never watch any games, students of the tournament may vigorously debate a nominee’s worth without reading a single opinion. The tournament would make it less, not more, likely that discussions of the desirability of a particular nominee would be based on a review of the judge’s actual work product.

III. SHOULDN’T WE COMPARE APPLES TO APPLES AND ORANGES TO ORANGES?

As other contributors to this Symposium have observed, the qualities that allow a judge to excel on a circuit court are not necessarily those that engender success on the Supreme Court. It is not self-evident why productivity on the courts of appeals, citation count, and independence are the qualities we most need in a future Supreme Court Justice. Just to take one example, productivity may be a proxy for diligence, and we can agree that we generally want diligent Justices, but the Supreme Court’s docket has been shrinking over the past years from an average of about 170 cases to roughly ninety-five cases per year.⁶⁴ Is the productivity measure an effort to boost flag-

63. *See id.* at 28.

64. Mark Tushnet, *Foreword: The New Constitutional Order and the Chastening of Constitutional Aspiration*, 113 HARV. L. REV. 29, 68 (1999) (“At its peak, the Court heard argument in about 170 cases per Term, whereas during the past five years it has heard argument in fewer than one hundred cases each year.”); *see also* Hart v. Massanari, 266 F.3d 1155, 1177 n.34, 1178 n.37 (9th Cir. 2001) (noting that in the 1999 Term, the United States Supreme Court heard seventy-seven cases, or less than nine majority opinions per Justice,

ging productivity of the Court? Or should the public regard a Supreme Court appointment as a well-earned retirement after hard work on the circuit?

Professors Choi and Gulati have exhaustively studied which apple is best on the assumption that it will make the best orange. The winner of the tournament has himself acknowledged the possibility that greatness as a circuit judge does not necessarily assure greatness as a Justice.⁶⁵ This observation is not just a criticism of the tournament but instead a suggestion for another use to which it might be put. Rather than using the measure to predict which appellate court judge would make the best Supreme Court Justice, the Choi/Gulati metric might be useful in identifying whom should be nominated to the circuit courts in the future.

If a consensus existed that the tournament accurately measured the desired characteristics of circuit judges—a strong assumption—then it might be used to identify the traits that impart success to circuit judges. Why are the judges that succeed in the tournament so much better than the rest of the pack? Academics have studied and debated the relationship between a judge's personal characteristics and the outcomes of particular types of cases.⁶⁶ A similar exercise might be undertaken with the tournament. A judge's score in the tournament might be related to her individual characteristics. Such a study might reveal that the best judges are lapsed academics, erstwhile appellate practitioners, or former district judges. Or, perhaps only a former editor-in-chief of her law school's law review thrives on the bench. If so, such information would help the President and Senate identify and evaluate candidates for the appellate bench. Whether such a study would reveal a short set of qualities that assure success on the appellate bench is uncertain. Moreover, it is not clear that we want a bench comprised of judges satisfying a narrow judicial "profile." We have more democratic tradition in this regard than other countries whose judges rise through the civil service and undergo rigorous, uniform training. Democracy can be unruly, and as evidenced by our history, we have long assumed that the judiciary's membership should reflect a variety of life experiences. However,

while Ninth Circuit judges during 1999 heard an average of 450 cases, averaging twenty majority opinions and 130 unpublished dispositions).

65. Howard Bashman, *20 Questions for Circuit Judge Richard A. Posner of the U.S. Court of Appeals for the Seventh Circuit*, HOW APPEALING (Dec. 1, 2003), at http://legalaffairs.org/howappealing/20q/2003_12_01_20q-appellateblog_archive.html#107025481874565902 ("I also note that judges who do well on lower courts sometimes disappoint as Supreme Court Justices.")

66. For two of the more systematic studies, see Gregory C. Sisk et al., *Charting the Influences on the Judicial Mind: An Empirical Study of Judicial Reasoning*, 73 N.Y.U. L. REV. 1377 (1998); and Gregory C. Sisk et al., *Searching for the Soul of Judicial Decision-making: An Empirical Study of Religious Freedom Decisions*, 65 OHIO ST. L.J. 491 (2004).

such a study would at least compare apples to apples, rather than apples to oranges.

If interest instead lies in what makes a successful Supreme Court Justice, the appropriate subjects of study are the Justices themselves. The task of measuring performance of Justices on the Supreme Court is no small feat, as other contributors to this Symposium have discussed.⁶⁷ If a measure of a Justice's success on the Court existed, however, it could be related to his or her characteristics in a fashion similar to that described for circuit judges above. The results of the studies of circuit judges and Supreme Court Justices could be compared to determine whether the individual characteristics that bred success in one position did so in the other. Although this proposal is a wholly different research enterprise than Professors Choi and Gulati's, it would test directly their strong assumption that traits that make for a good circuit judge also make for a good Justice. It would force us to rethink the criteria: Do we believe that productivity, citation counts, and independence are the hallmarks of the great Justices? How do the great Justices fare in a "Tournament of Justices?"⁶⁸

The failure to test the key assumption of their model illustrates that the tournament, although it involves measurement, is not really social science. Ordinal rankings, like David Letterman's Top Ten lists, or numerical measurements, like baseball batting averages, are mere quantification. For the reasons described previously in this Essay, whether the tournament—despite its elaborateness—actually captures the phenomenon it purports to is uncertain. The use of a measurement whose validity is untested imposes three costs. First, as previously discussed, the measure may distort the incentives of the subjects who know that they are evaluated according to this metric. Second, judges who rank low in the tournament because of its imprecision are unfairly besmirched. Professors Choi and Gulati pull their punches by claiming that they cannot infer anything about the low-ranking judges.⁶⁹ If so, then perhaps only top finishers in the tournament should be announced. Third, the tournament also harms the public. Should litigants who appear before lower-ranking judges

67. See, e.g., Lee Epstein et al., *The Role of Qualifications in the Confirmation of Nominees to the U.S. Supreme Court*, 32 FLA. ST. U. L. REV. 1145 (2005); Daniel A. Farber, *Supreme Court Selection and Measures of Past Judicial Performance*, 32 FLA. ST. U. L. REV. 1175 (2005); Richard A. Posner, *Judicial Behavior and Performance: An Economic Approach*, 32 FLA. ST. U. L. REV. 1259 (2005); David C. Vladeck, *Keeping Score: The Utility of Empirical Measurements in Judicial Selection*, 32 FLA. ST. U. L. REV. 1415 (2005).

68. We know of some Justices who, by any measure of productivity, will not do well. See David P. Currie, *The Most Insignificant Justice: A Preliminary Inquiry*, 50 U. CHI. L. REV. 466 (1983); Frank H. Easterbrook, *The Most Insignificant Justice: Further Evidence*, 50 U. CHI. L. REV. 481 (1983).

69. See Choi & Gulati, *Empirical Ranking*, *supra* note 2, at 75-77.

conclude that they have received an inferior brand of justice? If Professors Choi and Gulati believe their tournament accurately measures the adequacy of judicial performance, are they suggesting that litigants appearing before lower-ranking judges did not receive due process?

IV. CONCLUSION

In complex human endeavors, attempts to quantify the dimensions of success are inherently appealing. Sports fans who keep game statistics are a prime example. But what makes rankings so enjoyable to discuss and analyze is that at some level fans know that the numbers never measure all the aspects of an athlete's play. Some intangible residual is left out, and this left-out portion might be the most important feature of an athlete's play. In judging, where outcomes are even less readily observed than sports, the gap between numerical measures and actual performance is greater. Professors Choi and Gulati's attempt at measurement is a valiant one, but to some extent, they have told us what we already knew: Judges Posner and Easterbrook are exceptional talents, and the rest of the appellate bench is hard to distinguish from one another. Their measures are imprecise ones, and were judges to compete along these dimensions, society might bear considerable costs. Studies that expand our knowledge of the work of the judiciary deserve applause, but Professors Choi and Gulati have not yet presented a convincing case that choosing Supreme Court nominees by tournament should replace selecting them by the Constitution's more political process.