

PULLING FROM THE RANKS?: REMARKS ON THE
PROPOSED USE OF AN OBJECTIVE JUDICIAL
RANKING SYSTEM TO GUIDE THE SUPREME
COURT APPOINTMENT PROCESS

HON. BRUCE M. SELYA*

I.	INTRODUCTION.....	1281
II.	OBJECTIVE MEASURES OF MERIT AND THE SUPREME COURT APPOINTMENT PROCESS.....	1283
	A. <i>Do We Need Objective Measures of Merit to Identify the Most Qualified Candidates?</i>	1283
	B. <i>Do Objective Measures of Merit Serve the Goal of Political Transparency?</i>	1285
	C. <i>Are Objective Measures of Judicial Performance Apt Indicators of Qualification to Serve on the Court?</i>	1287
	1. <i>The Productivity Index</i>	1287
	2. <i>The Quality Index</i>	1289
	3. <i>The Independence Index</i>	1290
	4. <i>Recapitulation</i>	1291
	D. <i>The Siphoning Effect of Employing Objective Measures of Merit as an Initial Screen</i>	1291
III.	THE EFFECTS OF OBJECTIVE MEASURES OF MERIT ON THE PERFORMANCE OF THE FEDERAL APPELLATE JUDICIARY	1292
IV.	THE INSTITUTIONAL EFFECTS OF OBJECTIVE MEASURES OF MERIT ON THE FEDERAL APPELLATE JUDICIARY.....	1295
	A. <i>The Effect of Competitive Ranking Systems on Intracircuit and Intercircuit Relations</i>	1295
	B. <i>Life at the Bottom and Judicial Legitimacy</i>	1296
V.	CONCLUDING REMARKS.....	1297

I. INTRODUCTION

Although politics and ideology long have played a role in federal judicial appointment and elevation,¹ that role has swelled in recent years. Ongoing Senate confirmation battles over federal court nominees and rampant speculation about potential retirements from the Rehnquist Court have morphed into mainstream news. One critique of this ongoing fascination with the appointment process is that it is fundamentally out of focus. The contemporary debate centers on predicting how a putative Justice might (or might not) tip the balance on hotbed political issues rather than on merit or qualification for judicial service. Consequently, the debate says very little about how

* Judge, United States Court of Appeals for the First Circuit; Adjunct Professor of Law, Boston College Law School; Adjunct Professor of Law, Boston University School of Law. B.A., Harvard College, 1955; J.D., Harvard Law School, 1958.

1. See Erwin Chemerinsky, *Ideology and the Selection of Federal Judges*, 36 U.C. DAVIS L. REV. 619, 620 (2003) (observing that despite the curiosity that “every generation has the sense that it is the first to uncover that ideology has a role in the judicial selection process,” in fact, “[e]very President in American history, to a greater or lesser extent, has chosen federal judges, in part, based on their ideology”).

those dimensions might be measured. Traditionally, the American Bar Association has produced a qualitative measure of merit (which it provides to the President, the Attorney General, and the Senate Judiciary Committee) that evaluates the integrity, professional competence, and judicial temperament of each federal judicial nominee and designates each nominee as well qualified, qualified, or not qualified.² This system, though valuable in some aspects, has limited utility because its “thumbs-up/thumbs-down” approach provides meager information about the relative merits of the nominees. Moreover, its legitimacy as a nonpartisan measure has come under attack by researchers who suggest that it could be a disguised political device.³

Two academics, Professors Choi and Gulati, have reacted to the perceived hyperpoliticization of the federal judicial appointment and elevation processes by calling for the development of “objective” measures for evaluating judicial talent. To this end, Professors Choi and Gulati have put forth a particularly provocative proposal for a “Tournament of Judges” (the “Tournament”)—an ongoing empirical contest among federal appellate judges that would purport to tabulate objective measures of judicial performance (such as opinion publication rates, citation rates, and frequency of dissent) and would rank judges according to their overall scores.⁴ In the authors’ construct, the highest-ranked judge would be offered up as the heir apparent to the next Supreme Court vacancy.⁵ Choi and Gulati apparently believe that their Tournament will produce one of two desirable results: either politics will take a back seat to merit or, failing that, a politically motivated nomination will no longer be able to masquerade as merit-based.⁶ Even short of Supreme Court appointment, they posit, the Tournament would infuse the federal appellate bench with an “otherwise absent external incentive” for excellence in performance.⁷ A ranking system would hold judges accountable to high per-

2. AM. BAR ASS’N, STANDING COMMITTEE ON FEDERAL JUDICIARY: WHAT IT IS AND HOW IT WORKS 1 n.1, 3, 8 (2005), available at [http://www.abanet.org/scfedjud/Federal_Judiciary%20\(2\).pdf](http://www.abanet.org/scfedjud/Federal_Judiciary%20(2).pdf). The ABA Standing Committee’s ratings on judicial nominees can be accessed at <http://www.abanet.org/scfedjud/ratings.html>.

3. See, e.g., James Lindgren, *Examining the American Bar Association’s Ratings of Nominees to the U.S. Courts of Appeals for Political Bias, 1989-2000*, 17 J.L. & POL. 1, 6 (2001) (asserting that “[e]xtensive data analysis [of the ABA ratings] revealed different patterns for evaluating Clinton and Bush appointees”).

4. Stephen Choi & Mitu Gulati, *A Tournament of Judges?*, 92 CAL. L. REV. 299, 303 (2004) [hereinafter Choi & Gulati, *Tournament*].

5. See *id.* For a follow-up article in which the Tournament’s architects run the numbers to show how such a competition would operate, see 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*].

6. See Choi & Gulati, *Tournament*, *supra* note 4, at 301.

7. *Id.* at 304.

formance standards because the potential reward of elevation and the reputational benefits of superstar status would motivate them to excel.⁸ Conversely, the risks of decreased peer respect and public embarrassment would ensure good work even from those judges who have no realistic prospect of elevation or who are indifferent to judicial celebrity.

On the surface, these dual objectives—merit-based elevation and increased incentive to perform—may have a patina of plausibility. But there is more here than meets the eye. This Essay examines whether these objectives are worth pursuing in the abstract and then considers whether objective measures are likely to produce the results bruted by the sponsors of the Tournament.⁹

II. OBJECTIVE MEASURES OF MERIT AND THE SUPREME COURT APPOINTMENT PROCESS

In the abstract, the task of developing objective measures of judicial performance seems straightforward. Achieving objective measures would assist in the selection of Supreme Court Justices both by providing standardized information about merit—valuable in itself for identifying the best candidates—and by gleaning the extent to which ideological alignment, diversity concerns, or other non-merit-based factors might drive a particular process of nomination and confirmation. It is, however, instructive to peek beneath the coverlet.

A. *Do We Need Objective Measures of Merit to Identify the Most Qualified Candidates?*

The primary function of objective measurement systems (such as the proposed Tournament) is to provide a standardized set of data in order to foster informed decisionmaking. The assumption underlying the perceived need for such data is that the informal, ad hoc methods currently employed to determine which candidates make the final cut are somehow deficient or that they lead to bad results. That assumption is somewhat puzzling. Although the existing process may be discursive and sometimes opaque, no one has made the case that it produces disagreeable outcomes (that is, that it results generally in the appointment or elevation of unqualified jurists). By any reasonable measure, the Article III judiciary comprises an array of talented men and women. This can only mean that there is an underlying merit-based quality control system at work. This system operates effectively, if somewhat obscurely, to ensure that, politics aside, success-

8. *Id.*

9. Though I focus on the Choi/Gulati proposal, I take that proposal to be representative of a broader bid to introduce empirical measures of judicial performance into the federal judicial system.

ful candidates have passed a certain threshold of merit. Thus, the real impetus behind objective measures cannot merely be a desire to ensure that candidates are qualified.

In fact, the absence of debate about the merits of nominees quite likely indicates that an informal but highly merit-conscious system of preliminary screening is already in place. To illustrate their point that decisionmakers eschew merit-based discussion, Choi and Gulati point to examples of political opponents talking past one another (such as the situation in which a proponent of a candidate asserts that the candidate is highly qualified and, instead of challenging that assertion head-on, the opponent responds with her ideological objections to the candidate's elevation).¹⁰

This claim that politics too often drowns out the merits fails to acknowledge that Presidents rarely submit nominees who would not fare well under an objective ranking system and that, as a result of that preliminary vetting, most nominees who come before the Senate easily meet or exceed any reasonable set of merit-based benchmarks. Were a candidate's qualifications miserable or even comparatively underwhelming, one would expect the candidate's foes to pounce on the fact of mediocrity rather than to engage in ideological polemics. In most cases, however, the fact of nomination is shorthand for the achievement of a behind-the-scenes consensus about objective qualification, such that public debate shifts almost immediately to more contentious issues.

Indeed, structural pressures all but guarantee that a President will nominate candidates who are highly qualified in terms of intelligence, experience, and skill. By vesting the nomination power solely in the hands of the President,¹¹ the Constitution concentrates accountability in a single individual. The caliber of the nominees will reflect directly upon the President. Alexander Hamilton wrote in *The Federalist No. 76* that "[t]he sole and undivided responsibility of [the President] will naturally beget a livelier sense of duty and a more exact regard to reputation. [The President] will, on this account, feel himself under stronger obligations, and more interested to investigate with care the qualities requisite to the stations to be filled"¹² In contrasting presidential appointment with the alternative of appointment by a multimember assembly, Hamilton observed that in the latter case "the intrinsic merit of the candidate will be too often out of sight."¹³

10. See Choi & Gulati, *Empirical Ranking*, *supra* note 5, at 26 & n.3, 27; Choi & Gulati, *Tournament*, *supra* note 4, at 321 & n.55.

11. U.S. CONST. art. II, § 2, cl. 2.

12. THE FEDERALIST NO. 76, at 492 (Alexander Hamilton) (The Modern Library ed. 1937).

13. *Id.* at 493.

Nor is there a mandate or even an implicit norm that requires a President to offer seats on the Supreme Court as a reward for being the best in terms of intelligence, skill, or service on a lower court. Therefore, respect for the nominating power and regard for the many nuanced components that enter into the selection process counsel against lodging merit-based objections unless a particular nominee is in fact poorly qualified. Merit operates only as a threshold, not as the ultimate determinant. One can argue for increasing the height of the threshold limit, but even if that is done, there still will remain a group of aspirants who can exceed it. It is nothing short of utopian to think there will be only one.

In sum, for all that can be said about the lack of mainstream discussion of merit, there exists no plausible basis for a substantiated claim that the present process fails to yield top-notch Justices. If an objective measurement system is redundant because it identifies roughly the same pool of candidates that historically have been considered and chosen, there will be very little tolerance for any costs that such an innovation imposes. The question, then, is whether the game is worth the candle.

B. Do Objective Measures of Merit Serve the Goal of Political Transparency?

The idea of developing objective measures of merit to direct Supreme Court appointment grows out of a desire to reduce the politicization that has increasingly plagued the process of nominating and confirming appeals court judges (and that threatens to embroil future Supreme Court nominees). This is a matter of preference, not of constitutional mandate. The Constitution assigns the responsibility of choosing Supreme Court Justices to the political branches of the federal government,¹⁴ and therefore it is unsurprising that ideology factors into the exercise of the nomination and confirmation powers. As a normative matter, that is as it should be.¹⁵ Even assuming that the extirpation of politics from the nomination and confirmation processes is a goal supportable on other grounds—and that is a stretch—I have no faith that any system of rankings would achieve it. From my perspective, political transparency—and not some false promise of liberation from politics—is the touchstone in considering the value of objective measures.

Seen in this light, the argument for the use of objective measures is that while we may not be able to eliminate the politics, we can at

14. U.S. CONST. art. II, § 2, cl. 2.

15. See Chemerinsky, *supra* note 1, at 628 (arguing that “ideology should be considered because the judicial selection process is the key majoritarian check on an anti-majoritarian institution”).

least unmask the political subtexts that underlie conflicting claims that a particular aspirant is either an intellectual giant or an intellectual pygmy of the kind whom we consider to be qualified/not qualified to serve on the Supreme Court. The example that Choi and Gulati offer is a presidential nominee who ranks forty-second out of 160 active circuit court judges on the merit-based scale.¹⁶ In that situation, they assert, a neutral observer can conclude that some factor other than objective merit (say, ideology) is driving the selection.¹⁷

That is hardly rocket science. Yet another equally likely scenario demonstrates how a ranking system may serve to provide an impenetrable cover for an essentially ideological choice. While a President may have some explaining to do if he nominates number forty-two, he will escape the burden of revealing his calculus if he picks a candidate from the top tier (say, from the fifteen top-ranked circuit judges), regardless of whether he bases that selection on the candidate's ideological alignment. In that way, a President will be able to hide behind the very metric that the Tournament's advocates have intended as a means of making his motivations transparent. Although Choi and Gulati submit that the "introduction of a norm to apply objective criteria will force politicians to provide more justification for their selection,"¹⁸ it may very well accomplish exactly the opposite result.

The likely proliferation of multiple "objective" metrics also may undermine the goal of political transparency. Merit, like beauty, often lies in the eye of the beholder. If decisionmakers do come to rely on a merit-based ranking system, the lure of potential influence on the appointment process doubtless will spawn a market of competing objective indices, each claiming greater accuracy. Savants, docents, and other interested parties will lose little time in developing formulas designed to reach particular results. Consequently, any ranking system will face constant criticism that it is a proxy for either political affiliation or ideological leanings rather than for merit. I fear that the jousting between warring indices would repoliticize the selection process at a level even further removed from the relative qualifications of judicial aspirants. The goal of political transparency will not be served by a band of cleverly designed ranking systems that merely masquerade as objective measures.

16. Choi & Gulati, *Empirical Ranking*, *supra* note 5, at 28.

17. *Id.*

18. *Id.* at 30.

C. Are Objective Measures of Judicial Performance Apt Indicators of Qualification to Serve on the Court?

Even if one assumes that, in the abstract, an objective ranking system could assist in the selection of candidates for the Supreme Court and could improve the political transparency of the process, the fact remains that a ranking system is only as effective as its component criteria. Choi and Gulati posit that judicial merit can be measured by a composite of “productivity,” “quality,” and “independence.”¹⁹ To measure productivity, they factor in the number of majority and dissenting opinions a judge has written and the number of cases in which the judge has participated.²⁰ To gauge quality, they seize on what they believe to be an already accessible market test for that attribute: looking at how often and how prominently a judge’s opinions are cited by other courts and academics.²¹ Finally, to gauge judicial independence, they reason that disagreement, particularly opposition to the opinion of a colleague appointed by a President of the same party, indicates that the judge has a mind of her own, and thus dissent rates constitute a valid proxy for this attribute.²²

While there is arguably some relationship between these three indices and judicial merit, I doubt that the correlation is close enough to justify placing much weight on a composite ranking, let alone to justify using such a ranking as the primary filter for narrowing the field of candidates for elevation.²³ My concerns fall into two general categories. First, I sense a series of disjunctures between the stated criteria and the concept of judicial merit. Second, even if the proxies were acceptable, they would be inherently manipulable by the contestants. I consider each index in turn.

1. The Productivity Index

In evaluating the productivity index, my immediate concern is that this measure is underinclusive because it ignores a range of appellate activities. Though writing opinions is the primary work of the appellate judge, that activity captures only part of the occupation. In

19. *Id.* at 42 (defining the terms productivity, quality, and independence for the purposes of their article).

20. For a more detailed description of this methodology, see *id.* at 42-47.

21. *Id.* at 48.

22. See *id.* at 61-67 (explaining the methodology for calculating the independence measure).

23. Professors Choi and Gulati readily acknowledge that the Tournament’s measures are flawed in many ways but dismiss these defects by downplaying what is in many ways the Tournament’s most attractive selling point—that it actually will tell us something about merit so that we can evaluate candidates on that basis. To do this, they inch away from the question whether the Tournament gets it right when it comes to identifying merit and instead insist that so long as the Tournament unmask political forces, it is a valuable tool. See *id.* at 35-36.

the interest of expediency, judges decide many cases by issuing judgments or memorandum orders, without full-dress opinions. Those devices are utilitarian, and their appropriate use should not be discouraged. Judges also regularly serve on duty panels that decide motions and other procedural matters—while it is not glamorous work, it is, nonetheless, necessary. At the request of the President or the Chief Justice, some judges undertake service on bodies such as the Sentencing Commission, the Judicial Panel on Multi-District Litigation, the Foreign Intelligence Surveillance Court, and committees of the Judicial Conference. This service often limits the number of cases that a judge hears in his own court. Finally, many judges are forced to restrict their sitting schedules in order to carry out the administrative burdens that are essential to the operation of the federal courts.

While the underinclusivity of the index is arguably correctable to some extent, it highlights the ease with which one can rig a ranking system to produce particular results by selectively including certain factors and excluding others. Moreover, this deficiency reminds us that ranking systems are scarcely masters of nuance; a particular candidate will fare well only if she fits the mold that the index constructs. Designing that index is a normative task that demands some agreement about what merit means. This enhances the likelihood that competing versions of merit—versions that are likely inseparable in certain ways from judicial philosophy (read “ideology”)—will produce competing indices, thereby clouding an observer’s ability to discern merit.

My next criticism is that the productivity factor is deceptively straightforward. By placing productivity on a pedestal, the ranking indulges an assumption that we want Supreme Court Justices who are docket-movers. That is a wholly untested notion of merit. Indeed, the index assumes that those who write more slowly or less often are inefficient, lazy, distracted, or just plain incompetent. That assumption is a canard: it unfairly demeans judges who subscribe to a philosophy of restraint or those who believe that publication for publication’s sake tends to confuse the law.²⁴ The point is that the index has its biases—and more troubling still, those biases may not be immediately apparent.

Manipulability is also a concern. Judges, to varying degrees, exercise control over their own numbers. If a judge wanted to increase her productivity score, she might work harder, take more cases, and produce more opinions. So long as this were done without sacrificing quality, this option would bring about a higher ranking while helping

24. See Hon. Bruce M. Selya, *Publish and Perish: The Fate of the Federal Appeals Judge in the Information Age*, 55 OHIO ST. L.J. 405, 407-08 (1994).

the administration of justice. But a judge might avail herself of other options to effect a score increase without accomplishing any systemic gains. She could, for example, sacrifice quality for quantity, dash off gratuitous concurrences or dissents, shy away from demanding cases in search of easy prey, or use a host of other tactics. These myriad opportunities for “strategic gaming” of the numbers would further debilitate the effectiveness of the index as a proxy for merit.

Last—but far from least—there are real costs to introducing new behavioral incentives into an institution in which justice has long been thought the cardinal goal. Hasty and unnecessary opinions generally make for bad law, and manipulative behavior may erode both the credibility of the system and the collegiality that is so necessary to the effective functioning of an appellate court. The point is that even though a productivity index may incentivize genuine improvements, there is an equal or greater risk that it will encourage behaviors inimical to the development of decisional law and to the legitimacy of the federal courts.

2. *The Quality Index*

The assertion that consumption indicates quality may have some truth in it because judges and academics concerned with producing coherent rationales are unlikely to rely on poorly reasoned ones. At least two distortions, however, will likely interfere with the quality signal.

The first problem with culling a quality score from citation rates is that all citations are not created equal. A citation may signal that the authoring judge produced an apt statement of a commonly used standard or a path-breaking approach to a complex area of the law. Absent a sophisticated coding system—one that would doubtless involve subjective judgments—the rankings would not record such nuances.

A second problem (and perhaps one that we should associate more generally with all supposedly objective measures of judicial performance) arises from the fact that subsequent authors will come to rely upon the rankings themselves, particularly the quality index, as a shorthand for quality. This will create a sort of self-fulfilling prophecy. Judges who rank high on the quality index will likely earn respect in the legal community, thereby increasing the probability that other judges and academics will cite their opinions. Future citation rates will then say less about the quality of a cited opinion and more about the entrenched reputation of the original writing judge. This illustrates the Achilles’ heel of the Tournament proposal. Rankings tend to take on a life of their own, and many people tend to rely on them without knowing much about what they really signify. Put

bluntly, rankings too often disengage “the thing” from “the thing signified” and thereby frustrate the objective (informed, substantive deliberation) for which they purport to supply a foundation.

Moreover, the quality index is ripe for manipulation. Any judge worth his salt will tell you that there are ways to write opinions that make citation more likely. Judges can boost citation rates by writing longer opinions, publishing opinions that would otherwise go unpublished, eschewing quotations, taking controversial positions, or reaching for novel issues at the margin of a case. Accordingly, frequency of citation sometimes may signal better strategy rather than better quality.

3. *The Independence Index*

The independence index aims to capture the intellectual independence of the judge by measuring his willingness to disagree with his colleagues.²⁵ This is pure fiction: rates of dissent are an invalid proxy for judicial independence. Many judges write fewer dissents because they ascribe a relatively high value to the institutional good of courts speaking with a single voice, and they are willing to work toward developing a template to which the entire panel can subscribe. A low dissent rate for such a judge is a badge of honor, not of shame. In all events, it tells us nothing about his independence.

Even if a judge’s rate of dissent provides some tiny amount of information about her independence, it is not clear that independence in that sense is a desirable trait. After all, there plainly is a point at which dissenters cross the line from enriching thought into either intellectual preening or obstructionist polemicism. The sponsors of this proxy thus overlook the obvious danger in encouraging dissent for its own sake (or more precisely, for the sake of a better independence score).

Furthermore, Choi and Gulati have conceded that a particular ranking system could elect to place a negative weight on dissent.²⁶ This concession highlights another of the many points at which subjective choices—choices of judicial philosophy and ideology—make their entry into a supposedly objective measurement system. And, finally, the likelihood of manipulability is very high. Incentivized manipulation always troubles me—but it troubles me particularly in this context because it encourages dissent for the sake of dissent and, in the bargain, threatens both collegiality and the clarity of the law. The cruel irony of the independence index is that the act of measurement threatens to destroy the thing measured.

25. Choi & Gulati, *Empirical Ranking*, *supra* note 5, at 61-62.

26. *Id.* at 62.

I might add that this effect is far worse than the perverse incentive that Choi and Gulati attribute to the current appointment system.²⁷ They asseverate that judges now use voting, opinion writing, and dissent to signal their willingness, if elevated, to pander to the ideology of the appointing administration.²⁸ In my experience, this seldom occurs.

4. *Recapitulation*

The bottom line is that I cannot credit the claim that these three indices constitute objective measures that aptly gauge generally accepted notions of judicial merit.²⁹ Moreover, the costs of the Tournament proposal are yet to be calculated—and those costs must be justified against whatever meager benefits an actual Tournament might provide. I now turn to those costs.

D. *The Siphoning Effect of Employing Objective Measures of Merit as an Initial Screen*

The federal appellate judiciary is not a densely populated institution. Having narrowed their field to include only active federal appellate judges, Choi and Gulati place approximately 160 candidates in their Tournament.³⁰ Although they frankly acknowledge that there are normative and historical objections to choosing such an exclusive pool as the starting point for Supreme Court appointment³¹—after all, two of the nine sitting Justices come from other venues, and Justice Souter was a sitting federal appellate judge for only a day—they attempt a test run of the Tournament by further narrowing the sample to ninety-eight judges (eliminating latecomers and those who did not remain active through June of 2003).³²

Even more problematic than this apparent elitism is the fact that a tournament system aggrandizes attributes on the basis of ease of measurability rather than relevance to what really makes an ideal Supreme Court Justice.³³ Other dimensions, such as temperament, integrity, and worldliness, are left by the wayside.³⁴ Unless one actu-

27. *See id.* at 34.

28. *See id.* The authors concede that the threat of nomination blocking provides a check on overly explicit signaling; they contend, however, that judges engage in “stealth signal[ing]” to avoid proclaiming ideology so strongly that opinions and dissents arouse the other side to “muster its resources to block [the candidate].” *Id.*

29. Choi and Gulati have conceded that an actual Tournament system would not provide a perfect or even a nearly perfect measure of judicial merit. *See, e.g., id.* at 35-36; Choi & Gulati, *Tournament*, *supra* note 4, at 312.

30. *See* Choi & Gulati, *Empirical Ranking*, *supra* note 5, at 28.

31. *See id.* at 40; Choi & Gulati, *Tournament*, *supra* note 4, at 318-19.

32. Choi & Gulati, *Empirical Ranking*, *supra* note 5, at 40-41.

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

34. *See id.*

ally values publication, citation, and dissent rates as the most exalted of all attributes—and I doubt that any of us who are not brain-dead would commit to that proposition—one must acknowledge that a tournament system may skip over candidates who, considered holistically, are the most desirable.

In sum, the Tournament proposal aims to circumscribe the field of candidates and even to limit the bounds of discourse based on a bob-tailed version of merit. Self-imposed restraints of this order require heightened justification, and it does not seem to me that the elucidation of the role of politics in the selection process is sufficiently important to warrant an artificial scheme that uses a three-attribute measure of merit to tip the selection scales in favor of particular candidates. I do not place such great faith in these objective indices, nor have I discovered in the Tournament proposal any sustainable normative argument that rates of publication, citation, and dissent are the most important attributes of a Supreme Court Justice.

III. THE EFFECTS OF OBJECTIVE MEASURES OF MERIT ON THE PERFORMANCE OF THE FEDERAL APPELLATE JUDICIARY

Having explored the implications of employing objective measures of merit in the Supreme Court selection process, I now turn to a secondary consequence of a selection system that relies heavily on such measures: the incentive effect, that is, the influence that the ongoing operation of a ranking system that holds itself out to be merit-based will have on the everyday performance of the federal appellate judiciary. Choi and Gulati posit that promotional and reputational motivations will compel judges to modify their behavior in response to a ranking system.³⁵ They find this result to be desirable because they believe that the Tournament identifies excellence, and therefore the incentive to score high in the Tournament will extract optimal performance from federal judges.³⁶ I cannot quarrel with their assumption that judges will react to publication of these statistics. I am less comfortable, however, with the premise that these incentives will produce a better, fairer, and more efficient judicial system. In a democracy that relies heavily on a system of checks and balances, more accountability is generally welcomed. But the Tournament imposes accountability for the wrong actions—like a failed health care system that compensates physicians for cloning things but not for taking care of patients. It also threatens the judicial independence that, under our system of government, is constitutionally guaranteed. Thus,

35. See Choi & Gulati, *Tournament*, *supra* note 4, at 300, 313-14.

36. See *id.* at 304.

the Tournament's avowed mission to remedy "the lack of incentive to seek promotion"³⁷ invites scrutiny.

I begin with a brief review of the mechanism by which Choi and Gulati anticipate this transformation will occur. The basic idea is that federal appellate judges—on the whole, a group of high achievers (and, thus, competitive)—will conform their behavior to the criteria used in the Tournament in pursuit of promotion (or at least the bragging rights that accompany a favorable ranking).³⁸ The public, in turn, will benefit from the amount of attention and care the judges devote to the measured tasks. Even if we set aside all questions pertaining to (1) the competitiveness of judges and (2) the relationship between the proposed criteria and the development of admirable judicial traits, there remains the question whether the promised result justifies the introduction of an external pressure on judicial decisionmaking.

Choi and Gulati exert little effort in addressing the implications of an incentive system for judicial independence. They write:

Whatever other objections exist, the one that we do not see room for is the argument that the tournament would hurt judicial independence. If anything, the pressures that appellate judges may currently feel to attract political sponsors by making decisions that please those sponsors would be eliminated. Indeed, if there is an objection to our system at all, it is that judges will be made too independent under it. The tournament will thus have eliminated one of the few popular checks on an otherwise independent judiciary.³⁹

This is more a distraction than an answer to the question. Among other things, it disparages the emphasis that the Founders placed on the independence of the judiciary.⁴⁰ *The Federalist No. 78* heralds the "independent spirit in the judges which must be essential to the faithful performance of so arduous a duty."⁴¹ This spirit is embodied in Article III's tripartite guarantee of independence, which comprises lifetime appointment, undiminished compensation, and exclusive vesting of the judicial power of the United States.⁴²

Despite their importunings to the contrary, Choi and Gulati's Tournament places this independence in the crosshairs. By their own measure, the Tournament system is successful if it exerts a force on federal judges that causes them to modify their behavior in a way

37. *Id.* at 300.

38. *See id.* at 314-15.

39. *Id.* at 320-21.

40. THE FEDERALIST NO. 78, at 504-05 (Alexander Hamilton) (The Modern Library ed. 1937) ("The complete independence of the courts of justice is peculiarly essential in a limited Constitution.").

41. *Id.* at 508.

42. U.S. CONST. art. III.

that brings it more in line with the ranked criteria. They boast that this will provide a degree of accountability.

In this context, a certain measure of accountability is healthy. We already require federal judges to comply with ethical rules, subject them to impeachment for high crimes and misdemeanors, and allow Congress to control the federal courts' jurisdiction, budget, and rule-making authority. But Choi and Gulati's proposed method for ensuring stellar judicial performance is a horse of a different hue. It not only conflicts with the Founders' apparent distaste for perpetual job evaluation,⁴³ but it also has its roots in a desire for personal advancement. Under the Tournament model, the impulse that spurs judges to perform to the best of their abilities comes not from a desire to administer perfect justice but, rather, from a desire to pursue promotion or indulge in an ego trip. This is not the kind of accountability that we should aim to inculcate.

One response to the admonition that objective measurements will jeopardize judicial independence is of the "lesser evil" brand—servitude to the rankings will relieve judges of their existing servitude to politics, so that they no longer will feel the need to modify their behavior to suit the ideologies of a sitting President. The hope is that the primary incentive will be to score better on a merit index, so that the system will reward good performance rather than cynical politicking. The trouble with this argument is that it presumes—without an iota of proof—that the new incentive will eliminate the old one. What is far more likely, however, is that a judge attentive to her Supreme Court ambitions will be tempted to serve two gods: ideological purity and the Tournament rankings.

The second response to the judicial independence concern is that regardless of what motives a Tournament system arouses, one cannot call its incentives pernicious so long as the result—better judicial performance—is beneficial. According to this thesis, the end justifies the means: We should not care if judges serve themselves or serve justice; the only thing that matters is that they produce timely opinions, get cited, and register regular dissents.

With respect, that is smoke and mirrors. Motivation is material because it is connected to the quality of the judicial product and, consequently, to the strength of the combined indices as a proxy for merit. As pointed out earlier, incentive-based judicial products may be of lower quality and, because the Tournament system is not designed to detect lower quality products and treat them differentially, these incentive-based products quite probably will further erode the proxy value of the three criteria. Over time, judicial rankings will say

43. *See id.*

less about actual merit and more about agility—the ability to game the system.

The inevitable secondary effect of the Tournament system thus undermines its primary purpose. The initial impetus for the measurement system is the search for judges who exhibit certain qualities that comport with the ideal of a jurist who will serve justice well, but because this is a reward-offering search, a tournament may instead lead us to judges who, like Pavlov's dogs, respond well to incentives and who understand what manipulations will create an apparent match to that ideal.

IV. THE INSTITUTIONAL EFFECTS OF OBJECTIVE MEASURES OF MERIT ON THE FEDERAL APPELLATE JUDICIARY

If the President and the Senate were to embrace an empirical ranking system to assist them in their respective nomination and confirmation responsibilities, other consequences would follow. I focus here on the implications for relations within and between circuits and the potential impact on the legitimacy of the federal appellate judiciary.

A. *The Effect of Competitive Ranking Systems on Intracircuit and Intercircuit Relations*

It does not take a sophisticated analysis to point out that introducing an assessment system that pushes judges to vie for position against one another will create potentially unhealthy competitive pressures. Within circuits, the parade of horrors would look something like this: division of labor would become a constant bone of contention, particularly with respect to the assignment of opinions; the independence index would spawn superfluous dissents, thereby causing damage both to collegiality and to the rule of law; and the rankings would create an explicit pecking order for an already competitive process of hiring law clerks. In short, this sort of competition would have unhealthy consequences for the way that we work with one another.

A second set of consequences would imperil intercircuit relations and could well lead to the isolation of the law of each circuit. Notably, the particular quality index that Choi and Gulati have selected to plug into their composite formula primarily measures “outside citations,” meaning citations by judges from other courts.⁴⁴ Because the Tournament ascribes a high value to cross-circuit citation but no

44. Choi & Gulati, *Empirical Ranking*, *supra* note 5, at 71 (stating that the quality index consists of an adjusted figure representing the number of outside citations for a judge's top twenty opinions and the number of invocations).

value to intracircuit citation, a judge can cite to an intracircuit colleague without increasing her score but may be reluctant to cite an intercircuit colleague, who is also her competitor. After all, doing so would give the cited judge a boost in the rankings. This consequence may discourage citation to opinions from other circuits that, while not controlling authority, are illuminating and persuasive, even though such reliance would contribute to uniform development of federal law. In the worst-case scenario, this undesirable strain on the conversation between circuits as they encounter novel questions of law may even encourage artificial circuit splits or uncertainties about whether circuits agree or disagree on particular points of law.

Another consequence arises from the eventuality that the rankings will generate notions of prestige that attach to the reputation of the circuit at large. Circuits that do not host a high-ranking Tournament competitor may command less respect or attention from other judges, lawyers, academics, and clerkship aspirants. Those circuits may then suffer from circumscribed influence on the development of the law. Though effects of this kind are subtle, they are as difficult to control as they are to detect.

The more general point is that rankings tend to assert a peculiar power in our society, whereby they come to define our entire perception of the things they measure instead of providing a limited set of information about those things. And, in practical terms, it is impossible to cabin the influence of rankings to the narrow purpose for which they were designed.

B. Life at the Bottom and Judicial Legitimacy

Even though ranking systems are inherently relative, they inevitably give rise to labels that quickly become cemented to the name and reputation of each ranked member. If decisionmakers credit judicial rankings by relying upon them, the public and the legal community may come to perceive a judge ranked at or near the bottom not just as relatively less productive, less respected, or less independent than his peers but also as a rotten judge. This presumption of incompetence would emerge regardless of the fact that virtually the entire membership of the federal appellate bench is first-rate. In a ranked list, someone has got to wind up at the bottom.

The rankings also could warp the legitimacy of bottom-tier judges. Low rankings and their attendant consequences may alienate the judges who receive them. A system that effectively creates a second-class judiciary could come to define entire careers. I fear that an entrenched ranking system would make it far too easy to forget that every federal appellate judge has been entrusted by the President and the Senate to hold the federal judicial power for life.

V. CONCLUDING REMARKS

One thing can be said with certainty about Choi and Gulati's Tournament of Judges: it catapults us into an engaging metaphysical experiment. In the end, however, I think that it sends us down an unsightly path. The now-popular spectacle of university hierarchs scrambling to game the *U.S. News and World Report* rankings would caution us to avoid that path.

In my view, the judiciary would do well to keep its pikestaffs at parade rest and eschew the jousting that Choi and Galuti invite. Objective measures such as they describe are likely to supply information that is only marginally beneficial in a system that performs well in identifying talented candidates. That information is as likely to obscure political motives as it is to expose them. Moreover, the suggested measures may be wildly inexact proxies for merit. The gains that a Tournament would provide are modest at most—and the cost of them, should they materialize, is the infliction of substantial harm.

To cinch matters, artificially grafting the proposed measures of judicial performance onto the current selection process misconceives the concept of Supreme Court elevation by portraying a Court vacancy as the ultimate reward to which objectively deserving federal appellate judges are entitled. That is as wrong as wrong can be. We ought to understand nomination and confirmation as a complex search for an individual who will best serve the nation at a particular point in time. Suitability for service on the Court cannot be reduced to a matter of baseball card statistics. From that perspective, the effort to line up the federal judiciary into tidy rows, ordered exclusively by a handful of objectively measurable considerations, interferes with what is a much broader process. What is more, the business of the federal appellate judiciary is the administration of justice through the exercise of Article III jurisdiction. As it stands, a Supreme Court vacancy is an infrequent interruption of that business. Thus, it makes little sense to advertise the unlikely prospect that political lightning will strike as a legitimate daily preoccupation for federal appellate judges. It makes even less sense to hope that federal judges will indulge in such a distraction.

In the last analysis, federal appellate opinions are not applications for employment on the Supreme Court. Scoring judges as if that were the endgame takes too narrow a view of a President's prerogatives while at the same time encouraging the membership of an independent judiciary to subordinate judicial wisdom to the whims of personal ambition. Objectively speaking, the Tournament proposal strikes out on its pitch that it will produce merit-based decisionmaking and transparency.

