

# Duke Law Journal

---

VOLUME 58

APRIL 2009

NUMBER 7

---

## FOREWORD

### MEASURING JUDGES AND JUSTICE

JEFFREY M. CHEMERINSKY†

JONATHAN L. WILLIAMS††

Quantitative analysis has influenced and sometimes transformed many disciplines, including economics, finance, and public policy—even wine production, movie production, and athletic recruiting.<sup>1</sup> Its popularity varies, however, depending on whether its methods appear to produce results. For example, the “quants” who made portfolio management a science began to receive blame in 2008 for deflating the economy. Warren Buffett was particularly biting: “[B]eware of geeks bearing formulas.”<sup>2</sup> In sports, some teams have considered firing scouts who judge prospective players by watching them play and instead evaluating talent based on empirical evidence of

---

Copyright © 2009 by Jeffrey M. Chemerinsky and Jonathan L. Williams.

† Duke University School of Law, Class of 2009. Editor-in-Chief, *Duke Law Journal*, Volume 58.

†† Duke University School of Law, Class of 2009. Articles Editor, *Duke Law Journal*, Volume 58.

Thanks to Dean Erwin Chemerinsky, Professor Mitu Gulati, and Erin Blondel for their helpful comments.

1. IAN AYRES, *SUPER CRUNCHERS: HOW ANYTHING CAN BE PREDICTED* 1–9, 64–80, 144–49 (2007); Dennis Overbye, *They Tried to Outsmart Wall Street*, N.Y. TIMES, Mar. 10, 2009, at D1 (quoting Warren Buffet).

2. Overbye, *supra* note 1.

performance.<sup>3</sup> Empirical enthusiasts claim that statistics reveal valuable information that the plain eye and intuition overlook. But critics argue that the empirical movement shortchanges the value of traditional analytical methods.

Empirical scholars have begun to train these same tools on the judiciary. They have studied topics ranging from the economic effects of judicial systems to the influence of ideology on judicial decisionmaking. Unfortunately, empiricists have often failed to consult their studies' subjects—participants in the judicial system—about their research premises. At the same time, many in the legal system know little about this literature. Those few lawyers, legal scholars, and judges who are aware of empirical findings have often openly resisted them, particularly those studies that emphasize the attitudinal model, which, in its strongest form, claims that judges decide cases based on their policy preferences rather than legal doctrine. The attitudinal model, they say, betrays a cramped view of the law and represents a fundamental misunderstanding of what judges do.

Despite these criticisms, the still-nascent field of empirical legal studies is growing. As databases become more comprehensive and the technology necessary to process data improves, empirical scholars are likely to produce more and more sophisticated analyses of legal decisionmaking.<sup>4</sup> But what has the discipline offered the law so far? Many courts already use data-driven analyses to improve administrative functions, such as assigning cases, monitoring potential biases, and creating issue-specific courts. If quantitative approaches are valuable in these areas, could they also offer useful lessons for traditional legal analysis? In light of the legal community's objections that empiricists fail to understand how lawyers and judges really argue and decide cases, we thought it would be useful to bring empirical scholars and the people they study together for a day of discussion.

The conference focused on several questions:

1. How, if at all, can quantitative measurements explain judicial behavior?

---

3. See generally MICHAEL LEWIS, *MONEYBALL: THE ART OF WINNING AN UNFAIR GAME* (2003) (describing the rise of quantitative scouting in baseball); Michael Lewis, *The No-Stats All-Star*, N.Y. TIMES, Feb. 13, 2009, at MM26 (describing the same in basketball).

4. See AYRES, *supra* note 1, at 130–44 (detailing advances in empirical tools and studies).

2. What objective criteria for evaluating judicial decisionmaking would help judges improve?
3. Why do judges resist the attitudinal model and other models of judicial behavior?
4. How could researchers adjust their models or methods to better reflect how the legal system operates?

This Symposium emerged from Dean David Levi and Professor Mitu Gulati's course at Duke Law School on the study of judicial behavior. Dean Levi, a former federal district judge, and Professor Gulati, whose work includes empirical studies of the judiciary, helped us assemble a group of empirical scholars to present papers and professors and judges to respond. Trial and appellate judges came from federal and state courts from across the country. Many of them have held administrative roles within the judiciary in addition to their day-to-day courtroom duties. Some judges were familiar with the empirical literature; others were not. The same was true of the doctrinal scholars we invited. This mix, we hoped, would provide a range of perspectives to improve future empirical research and the legal community's familiarity with the empirical literature.

The day-long conference, held on February 6, 2009, did not disappoint. Sometimes disagreeing and often agreeing, our participants covered a variety of the literature's strengths and weaknesses. In the pages that follow, several issues recur.

First, what data are relevant? Dean Levi, in his review of Judge Posner's book, *How Judges Think*,<sup>5</sup> faults Posner for ignoring unpublished dispositions.<sup>6</sup> If one wants to know how judges think day to day, Levi argues, ignoring cases that do not make it into the reporters—the vast majority—does little to enhance understanding of how judges usually think.<sup>7</sup> In unpublished opinions, judges may approach cases differently. If one asks a slightly different question—what is the effect of a particular rule—one would want to see how judges apply the rule in all cases, including unpublished ones.<sup>8</sup> Using only reported cases provides a narrow, possibly misleading sample, whereas incorporating unpublished opinions would enable a lawyer to

---

5. RICHARD A. POSNER, *HOW JUDGES THINK* (2008).

6. David F. Levi, *Autocrat of the Armchair*, 58 DUKE L.J. 1791, 1801–04 (2009) (reviewing POSNER, *supra* note 5).

7. *Id.* at 1803.

8. *Id.*

make an empirical argument about how the rule actually works.<sup>9</sup> Consider *United States v. Leon*,<sup>10</sup> in which the Supreme Court refused to apply the exclusionary rule to evidence officers obtained unconstitutionally when the officers reasonably relied on an invalid warrant.<sup>11</sup> According to the Court, the exclusionary rule, which prevents future misconduct by threatening that courts will exclude illegally obtained evidence, was inappropriate because the officers did not know they were violating the Constitution. After all, the Court implied, officers will not adjust their behavior to conform to constitutional rules when they believe they are already following the law. In his concurrence, Justice Blackmun observed that the majority's decision rested on an empirical judgment about the effects of the exclusionary rule—a judgment the Court should be prepared to reconsider based on new empirical evidence. Unpublished dispositions would provide the data necessary to evaluate *Leon's* effects and whether the Court should reconsider that decision.<sup>12</sup>

Second, how should scholars design empirical studies of legal decisionmaking? For example, once a scholar decides what to test, how does the scholar weigh the data? Professors Brennan, Epstein, and Staudt suggest that the Supreme Court may have consciously aided the Roosevelt administration's economic program during the Great Depression.<sup>13</sup> Comparing the percentage of pro-government votes in tax cases during a routine downturn and during the Great Depression, the professors determine that the Court voted more often in favor of the government during the Great Depression than in other periods, whereas the Court tended to vote against the government during routine downturns and for the government during normal periods of economic growth.<sup>14</sup> Professor Young and Ms. Blondel respond that Brennan, Epstein, and Staudt's approach is misleading because looking only to the number of pro-government

---

9. *Id.*

10. *United States v. Leon*, 468 U.S. 897 (1984).

11. *Id.* at 900.

12. To a large extent, this disagreement over how to handle unpublished opinions in the datasets likely reflects the legal community's disagreement regarding which opinions courts should publish and what weight courts and lawyers should give unpublished opinions.

13. Thomas Brennan, Lee Epstein & Nancy Staudt, *Economic Trends and Judicial Outcomes: A Macrotheory of the Court*, 58 DUKE L.J. 1191, 1196 (2009).

14. *Id.* at 1221.

case outcomes ignores the relative significance of cases.<sup>15</sup> For example, in *United States v. Butler*,<sup>16</sup> the Supreme Court struck down a key piece of New Deal tax legislation,<sup>17</sup> which “in its practical significance . . . may outweigh five or ten ordinary victories in ordinary tax cases.”<sup>18</sup> Young and Blondel’s criticism implies that when assessing how the Court decides cases in an area of law, the most important variable is the net effect of the Court’s decisions, not the number of the Justices’ votes. This discussion about how to measure particular variables that may determine studies’ outcomes will continue to receive serious scrutiny.

A third and related issue is distinguishing empiricists’ descriptive and normative claims. For example, Professors Choi, Gulati, and Posner rank state supreme courts based on three measures: productivity, opinion quality, and independence.<sup>19</sup> To measure these subjective characteristics, they use the number of published opinions,<sup>20</sup> out-of-state citation numbers,<sup>21</sup> and how often a judge votes with opposite-party judges,<sup>22</sup> respectively. The authors concede that these proxies are imperfect.<sup>23</sup> Determining the right proxies involves, as Dean Chemerinsky points out, difficult normative choices.<sup>24</sup> Choosing one definition of quality judging necessarily excludes other important measures, and because empirical studies need quantifiable data, they may rely too heavily on measurable criteria and neglect important but hard-to-measure criteria. Chief

---

15. Ernest A. Young & Erin C. Blondel, Response, *Does the Supreme Court Follow the Economic Returns? A Response to A Macrotheory of the Court*, 58 DUKE L.J. 1759, 1771–33 (2009).

16. *United States v. Butler*, 297 U.S. 1 (1936).

17. *Id.* at 74.

18. Young & Blondel, *supra* note 15, at 1770.

19. Stephen J. Choi, Mitu Gulati & Eric A. Posner, *Judicial Evaluations and Information Forcing: Ranking State High Courts and Their Judges*, 58 DUKE L.J. 1313, 1320 (2009).

20. *Id.* at 1320.

21. *Id.* at 1321.

22. *Id.* at 1323.

23. *Id.* at 1314; see also Scott Baker, Adam Feibelman & William P. Marshall, Response, *The Continuing Search for a Meaningful Model of Judicial Rankings and Why It (Unfortunately) Matters*, 58 DUKE L.J. 1645, 1647 (2009) (“[W]e strongly suspect that the attributes that the authors select . . . constitute relatively minor aspects of judicial quality.”); Laura Denvir Stith, Response, *Just Because You Can Measure Something, Does It Really Count?*, 58 DUKE L.J. 1743, 1748 (2009) (criticizing the authors’ judicial-quality metric as encouraging a “numbers game” that emphasizes “quantity over quality”).

24. Erwin Chemerinsky, Response, *No Warrant for Radical Change: A Response to Professors George and Guthrie*, 58 DUKE L.J. 1691, 1700–01 (2009).

Justice Stith of Missouri observes that preferring measurable criteria can create bad incentives to improve in the rankings instead of improving real quality.<sup>25</sup> These difficulties, which the authors acknowledge,<sup>26</sup> will generate controversies over any decisions to use proxies, especially in sensitive areas such as judicial rankings.

Fourth, what role should empirical conclusions have in crafting policy? Professors George and Guthrie determine that the Supreme Court would reach the same decision in vast majority of cases if it sat in panels.<sup>27</sup> From this foundation, they argue that the Court should sit in panels so that it could correct lower courts' errors more often, leaving the en banc procedure available for contentious cases.<sup>28</sup> Judge Boudin<sup>29</sup> and Dean Chemerinsky<sup>30</sup> accept their premise, but they reject George and Guthrie's proposal. Boudin argues that the Court's primary function is not, and should not be, error correction.<sup>31</sup> Circuit splits are not a significant problem; far more important is that the Justices have time to decide cases of national importance and serve as ambassadors for the judiciary. Chemerinsky argues that the Court's en banc hearings are so ingrained that a Court sitting in panels would lack legitimacy.<sup>32</sup> Additionally, he points out that it is difficult to determine which cases actually present errors the Court needs to correct, suggesting that the problem George and Guthrie seek to solve may simply be a phantom.<sup>33</sup>

Finally, the legal community has greeted many empirical findings skeptically. Professor Richman suggests that legal professionals may simply feel uncomfortable embracing unfamiliar empirical analyses by nonlawyers who disregard more traditional forms of legal reasoning.<sup>34</sup> Richman attributes the resistance to empirical evaluations of judicial behavior in part to the legal community's strong sense of

---

25. Stith, *supra* note 23, at 1748.

26. Choi et al., *supra* note 19, at 1313 (acknowledging that the measures are "coarse" (italics omitted)).

27. Tracey E. George & Chris Guthrie, *Remaking the United States Supreme Court in the Courts' of Appeals Image*, 58 DUKE L.J. 1439, 1462–65 (2009).

28. *Id.* at 1453–68.

29. Michael Boudin, Response, *A Response to Professors George and Guthrie, Remaking the United States Supreme Court in the Courts' of Appeals Image*, 58 DUKE L.J. 1685 (2009).

30. Chemerinsky, *supra* note 24.

31. Boudin, *supra* note 29, at 1685–86.

32. Chemerinsky, *supra* note 24, at 1698–700.

33. *See id.* at 1694 ("But to speak of 'errors' is to beg enormously difficult questions as to what is 'correct' as opposed to 'erroneous.'").

34. Barak Richman, Response, *On Doctors and Judges*, 58 DUKE L.J. 1731, 1740 (2009).

professionalism, which may cause lawyers to discount outsiders' evaluations merely because they are not members of the community.<sup>35</sup> On the other hand, some in the legal community object that empiricists overlook how conventional legal explanations can inform empiricists' hypotheses or explain their findings.<sup>36</sup> Many participants also suggest that how empiricists approach and describe their studies may contribute to mutual wariness. In particular, empiricists frequently charge that judges make decisions based on "politics,"<sup>37</sup> which unfortunately resonates with popular debates about judicial activism and conflates judges' long-acknowledged discretion when deciding difficult cases with "political" decisionmaking.<sup>38</sup> Judge Boudin and Dean Levi therefore ask scholars to use the term "politics" more carefully.<sup>39</sup> In addition, any claim that judges operate beyond the bounds of law is a serious one, and the legal community is likely to react negatively.<sup>40</sup>

We convened the *Duke Law Journal* Conference on Measuring Judges and Justice so that empirical scholars and legal professionals could discuss these issues. All of the articles presented at the conference follow, and some of the judges and doctrinal legal scholars also chose to publish their responses. We hope this issue provides an overview of the contributions that empirical scholars have tried to make and some criticisms of their efforts. To the extent this Symposium succeeds, we have our participants to thank.

---

35. *See id.* at 1740 ("[J]udges and legal academics often exhibit disregard for, and often hostility toward, outsiders who employ nonlegal analytical methods to assess judicial quality, understand judicial reasoning, or predict judicial outcomes.").

36. *See, e.g.,* Young & Blondel, *supra* note 15, at 1772 (suggesting that "conventional legal explanations may well resolve" some of Professors Brennan, Epstein, and Staudt's findings).

37. *See, e.g.,* J. Mark Ramseyer, *Predicting Judicial Outcomes Through Political Preferences: The Japanese Supreme Court and the Chaos of 1993*, 58 DUKE L.J. 1557, 1558–59 (2009) (arguing that judges "indulge their political biases" from the bench, but that their ability to do so is an indicator of judicial independence).

38. H. Jefferson Powell, Response, *A Response to Professor Knight, Are Empiricists Asking the Right Questions About Judicial Decisionmaking?*, 58 DUKE L.J. 1725, 1725–26 (2009).

39. Boudin, *supra* note 28, at 1689 ("Figuring out why judges decide cases the way they do is a worthy enterprise; not so scoring judicial results as 'political.'"); Levi, *supra* note 6, at 1795 n.15 (calling Judge Posner's use of the word "political" "unfortunate").

40. *See, e.g.,* Baker et al., *supra* note 23, at 1641 ("[I]t is troublesome to think that [Justices would] decid[e] cases based on economic conditions unrelated to the merits of the legal issue before them. They have no constitutional authority to act in this manner.").

We thank the participants in the *Duke Law Journal* Conference on Measuring Judges and Justice:

- Scott Baker, University of North Carolina School of Law;
- Larry Baum, The Ohio State University Political Science Department;
- Judge Michael Boudin, United States Court of Appeals for the First Circuit;
- Thomas Brennan, Northwestern University School of Law;
- James Brudney, The Ohio State University Moritz College of Law;
- Dean Erwin Chemerinsky, University of California, Irvine School of Law;
- Stephen Choi, New York University School of Law;
- Judge Allyson Duncan, United States Court of Appeals for the Fourth Circuit;
- Lee Epstein, Northwestern University School of Law;
- Tracey E. George, Vanderbilt University Law School;
- Mitu Gulati, Duke University School of Law;
- Chris Guthrie, Vanderbilt University Law School;
- Chief Judge Robert Henry, United States Court of Appeals for the Tenth Circuit;
- Justice Robin E. Hudson, Supreme Court of North Carolina;
- Jack Knight, Duke University School of Law and Political Science Department;
- Dean David F. Levi, Duke University School of Law;
- Judge Gerard E. Lynch, United States District Court for the Southern District of New York;
- Kevin T. McGuire, The University of North Carolina at Chapel Hill Political Science Department;
- William P. Marshall, University of North Carolina School of Law;
- Chief Justice Ruth V. McGregor, Arizona Supreme Court;
- Eric A. Posner, The University of Chicago Law School;
- H. Jefferson Powell, Duke University School of Law;
- J. Mark Ramseyer, Harvard Law School;
- President Tom Ross, Davidson College and former director of the North Carolina Administrative Office of the Courts;
- Barak D. Richman, Duke University School of Law;
- Judge Lee H. Rosenthal, United States District Court for the Southern District of Texas;
- Judge Barbara J. Rothstein, United States District Court of the Western District of Washington and director of the Federal Judicial Center;
- Chief Judge Anthony J. Scirica, United States Court of Appeals for the Third Circuit;
- Christopher H. Schroeder, Duke University School of Law;
- Joanna M. Shepherd, Emory Law School;
- Nancy Staudt, Northwestern University School of Law;
- Chief Justice Laura Denvir Stith, Supreme Court of Missouri;
- Ernest A. Young, Duke University School of Law;
- Judge John Shepard Wiley Jr., Los Angeles Superior Court.