

HARVARD LAW REVIEW

© 2011 by The Harvard Law Review Association

ARTICLES

LIGHTENED SCRUTINY

Bert I. Huang

TABLE OF CONTENTS

INTRODUCTION	1111
I. DEFERENCE ADRIPT?.....	1116
A. <i>The Judges' Hypothesis</i>	1118
B. <i>In Search of Evidence</i>	1119
II. A NATURAL EXPERIMENT: "THE SURGE"	1121
A. <i>The Unusual Origins of the Surge</i>	1122
B. <i>Toward a Causal Story</i>	1123
C. <i>A Second Experiment</i>	1126
III. FINDINGS: LIGHTENED SCRUTINY	1127
A. <i>The Data</i>	1127
B. <i>Revealed Deference</i>	1130
C. <i>Alternative Stories?</i>	1133
IV. PROBLEMS FOR THEORY AND POLICY	1137
A. <i>"Silent Splits"</i>	1139
B. <i>Reimagining Boundaries</i>	1142
C. <i>Detecting the Deference</i>	1144
V. CONCLUSION: AN INTEGRITY DILEMMA?	1145
APPENDIX	1147

LIGHTENED SCRUTINY

Bert I. Huang*

The current anxiety over judicial vacancies is not new. For decades, judges and scholars have debated the difficulties of having too few judges for too many cases in the federal courts. At risk, it is said, are cherished and important process values. Often left unsaid is a further possibility: that not only process, but also the outcomes of cases, might be at stake. This Article advances the conversation by illustrating how judicial overload might entail sacrifices of first-order importance.

I present here empirical evidence suggesting a causal link between judicial burdens and the outcomes of appeals. Starting in 2002, a surge of cases from a single federal agency flooded into the circuit courts. Two circuits bore the brunt, with their caseloads jumping more than forty percent. The other circuits were barely touched, by comparison. To sort cause from effect, I focus on outcomes not in the surging agency cases, but instead in a separate category: civil appeals. The two circuits flooded with agency cases began to overrule district court decisions less often — in the civil cases. This evidence of evolving deference raises the possibility of “silent splits”: divergences among the circuits in their levels of appellate scrutiny, due not to articulated disagreements but to variation in caseloads.

* Associate Professor of Law, Columbia Law School. I wish to thank Robert Akerlof, Oren Bar-Gill, Michael Boudin, Sam Bray, Richard Briffault, José Cabranes, Glenn Cohen, Rosalind Dixon, Ben Edelman, Noah Feldman, Sydney Foster, Katherine Franke, Jeanne Fromer, Abbe Gluck, Claudia Goldin, Jeff Gordon, Lani Guinier, Scott Hemphill, Christine Jolls, Louis Kaplow, Larry Katz, Robert Katzmann, Martin Kurzweil, Joe Landau, Maggie Lemos, Katerina Linos, Debra Livingston, Gerard Lynch, Wendy Martinek, Gillian Metzger, Henry Monaghan, Ed Morrison, Trevor Morrison, Jon Newman, J.J. Prescott, Alex Raskolnikov, David Rosenberg, Bertall Ross, Robert Sack, Dan Schwarcz, Robert Scott, Steve Shavell, Micah Smith, Holger Spemann, Matt Stephenson, Kate Stith, Lior Strahilevitz, Jeannie Suk, Neal Ubriani, Catherine Wolfe, Mark Wu, Kathy Zeiler, and the editors of this journal for their insights and comments on drafts. Workshop audiences at Columbia, Georgetown, Harvard, University of Chicago, Vanderbilt, and Yale, and at ALEA, CELS, and NBER conferences, as well as friends and colleagues, offered helpful suggestions. I thank Lena Husani and Caitlin Street for excellent research assistance and Tom Miles for sharing data. The Olin Center at Harvard; the Harvard Project on Justice, Welfare & Economics; the Paul and Daisy Soros Fellowship; and an NSF graduate research grant offered financial support. The views expressed are my own.

INTRODUCTION

Arbitrary and capricious. Abuse of discretion. De novo. Clear error. Emphatic phrases like these are what our judiciary uses to mind the minders — to regulate how the appeals courts in turn review the rulings of the trial courts.¹ Announced in nearly every circuit court opinion,² such standards of review are doctrines that reflect a received wisdom about these courts' relative institutional strengths.³ They are both artifacts and devices of our lawmakers' efforts to fine-tune the sharing of legal authority within the judiciary.⁴

Continuing contests and circuit splits over such doctrines reveal the real stakes for both litigants and the courts in defining these relations of review.⁵ Once set, the standards are embedded in the common law.⁶ The point of enshrining them in precedent, of course, is to promote uniformity and stability. After all, it seems only fair that the same type of issue on appeal be reviewed with the same degree of scrutiny from case to case, from place to place, and from one year to the next.

But for all this creating and citing of standards of review, how much consistency is being ensured, in fact? Judges themselves have raised doubts, and their doubts arise from a singular source: a half century of unrelenting growth in judicial workload. Judge Richard Posner, for one, sensed that “one consequence of the heavy caseload pressures on the courts of appeals has been an increase in the defer-

¹ Consider the familiar (if often fictional) binary of law versus fact: great deference is said to be due to a trial court's finding of fact; a ruling of law invites full reconsideration under de novo review, for which “no form of appellate deference is acceptable.” *Salve Regina Coll. v. Russell*, 499 U.S. 225, 238 (1991); *see also Bose Corp. v. Consumers Union of U.S., Inc.*, 466 U.S. 485, 498–501 (1984) (describing deferential standard of review for findings of fact generally, but not for “First Amendment questions of ‘constitutional fact,’” *id.* at 509 n.27).

² Moreover, every appellant's brief is required to state the proper standard of review, for each issue raised, under the current Federal Rules of Appellate Procedure. *See* FED. R. APP. P. 28(a)(9)(B).

³ *See Salve Regina Coll.*, 499 U.S. at 233; *see also Miller v. Fenton*, 474 U.S. 104, 114 (1985); Henry P. Monaghan, *Constitutional Fact Review*, 85 COLUM. L. REV. 229, 237 (1985).

⁴ *See Monaghan, supra* note 3, at 234–39.

⁵ *See Concrete Works of Colo., Inc. v. City & Cnty. of Denver*, 540 U.S. 1027, 1033 (2003) (Scalia, J., dissenting from denial of certiorari) (“[T]he case is worthy of the Court's review because it presents a clear Circuit split on the standard of appellate review for the ‘strong basis in evidence’ requirement.”); *see also, e.g., Gall v. United States*, 128 S. Ct. 586 (2007) (directing use of deferential abuse-of-discretion standard of review for federal criminal sentences); *Rita v. United States*, 127 S. Ct. 2456, 2462 (2007) (resolving circuit split over presumption-of-reasonableness review of federal sentences falling within Guidelines-recommended ranges); *Cooper Indus., Inc. v. Leatherman Tool Grp., Inc.*, 532 U.S. 424, 431 (2001) (resolving split over standard of review for constitutionality of punitive damages awards); *Ornelas v. United States*, 517 U.S. 690, 695, 699 (1996) (resolving split over standard of review for findings of reasonable suspicion to stop and probable cause to make a warrantless search).

⁶ This process includes glosses on standards set out in rules and statutes, such as in FED. R. CIV. P. 52(a) and in the Administrative Procedure Act, 5 U.S.C. § 706(2)(A) (2006).

ence paid by those courts to the rulings made by district judges.”⁷ Similarly, pointing to “troubling” data showing a long-term decline in the rates of appellate reversals, with caseloads growing all the while, Judge John Gibbons wondered if the circuit courts’ “remarkable achievement in productivity has been attained at least in part by the adoption of a posture of increased deference to the rulings of the courts we’re supposed to be supervising.”⁸

These judges’ hypothesis, in essence, is that but for heavy caseloads, some appeals would have succeeded that in fact did not. Who wins or who loses a given case could turn on how many other cases are in the court’s queue. Such a hypothesis deserves, but has lacked, a serious test. It is not that the caseload problem has gone unnoticed; to the contrary, there is a flourishing “crisis of volume” literature, itself voluminous.⁹ But those debates have mostly addressed process values, or the mechanics of appeals, rather than their outcomes.¹⁰ The litera-

⁷ RICHARD A. POSNER, *THE FEDERAL COURTS: CHALLENGE AND REFORM* 345 (1996). Judge Posner explains that “[a] possible consequence is that fewer errors made by district courts are being corrected — an example of an undesired by-product of the growth in the caseload.” *Id.* at 176. He also identifies a similar effect in the district courts, observing that the “least visible but probably most important way in which the pressure of a growing caseload had resulted in streamlining or corner cutting” is the “sub rosa redefinition [by district courts] of the standards for granting summary judgment and for dismissing a complaint for failure to state a claim.” *Id.* at 178.

⁸ John J. Gibbons, *Maintaining Effective Procedures in the Federal Appellate Courts*, in *THE FEDERAL APPELLATE JUDICIARY IN THE TWENTY-FIRST CENTURY* 22, 23 (Cynthia Harrison & Russell R. Wheeler eds., 1989) [hereinafter Gibbons, *Maintaining Effective Procedures*]; see also John J. Gibbons, *Illuminating the Invisible Court of Appeals*, 19 *SETON HALL L. REV.* 484 (1989).

⁹ This literature, too, goes back at least half a century, having engaged leading scholars and judges of a prior generation. See, e.g., Henry J. Friendly, *Averting the Flood by Lessening the Flow*, 59 *CORNELL L. REV.* 634, 642 (1974); Ruth Bader Ginsburg, *Reflections on the Independence, Good Behavior, and Workload of Federal Judges*, 55 *U. COLO. L. REV.* 1, 7–13 (1983); Lewis F. Powell, Jr., *Are the Federal Courts Becoming Bureaucracies?*, 68 *A.B.A. J.* 1370, 1371 (1982); Charles Alan Wright, *The Overloaded Fifth Circuit: A Crisis in Judicial Administration*, 42 *TEX. L. REV.* 949, 949 (1964). In fact, the phrase “crisis of volume” seems to have come into use in the 1970s. See, e.g., DANIEL J. MEADOR, *APPELLATE COURTS: STAFF AND PROCESS IN THE CRISIS OF VOLUME* (1974).

¹⁰ See, e.g., Mitu Gulati & C.M.A. McCauliff, *On Not Making Law*, *LAW & CONTEMP. PROBS.*, Summer 1998, at 157, 159–60 (focusing on “shortcuts” circuit courts take to manage their workloads, such as short-form opinions); Richard A. Posner, *Will the Federal Courts of Appeals Survive Until 1984? An Essay on Delegation and Specialization of the Judicial Function*, 56 *S. CAL. L. REV.* 761, 767–77 (1983) (describing consequences of the “delegation of the judicial function to law clerks and staff attorneys,” *id.* at 774); William M. Richman & William L. Reynolds, *Elitism, Expediency, and the New Certiorari: Requiem for the Learned Hand Tradition*, 81 *CORNELL L. REV.* 273, 284 (1996); Lauren K. Robel, *The Myth of the Disposable Opinion: Unpublished Opinions and Government Litigants in the United States Courts of Appeals*, 87 *MICH. L. REV.* 940, 940–41 (1989).

There is, of course, not always a sharp line between process and outcomes. See, e.g., Eric Helland & Jonathan Klick, *The Effect of Judicial Expedience on Attorney Fees in Class Actions*, 36 *J. LEGAL STUD.* 171 (2007) (describing the effect of court congestion on fee awards); Lee Epstein, William M. Landes & Richard A. Posner, *Why (and When) Judges Dissent: A Theoretical*

ture has lamented (or celebrated) the erosion of traditional practices, focusing on the decline of oral argument, say, or the increasingly routine use of summary orders in lieu of published opinions. By contrast, save for a few exceptions,¹¹ it has underplayed what to some must seem more pressing concerns: for the parties bringing appeals, their chances of winning; and for the designers of legal policy, coherence and order in the sharing of judicial power.

Yet as far as it goes, the judges' hypothesis stops short. It leaves unsaid a subtle but direct implication. To say that docket pressure can alter the nature of appellate scrutiny is to allow the possibility of variation not only across time, but also across jurisdictions. It is to recognize that circuit "splits," in a sense, may appear in the de facto intensity of review, arising not from reasoned disagreement but from arbitrary variation in caseloads. The same case on appeal might come out differently in one circuit than in another for a reason having little to do with the law of each circuit — that reason being a "spillover effect" from the presence of other, unrelated cases.

This Article attempts to isolate the impact of judicial burdens on the outcomes of appeals in a systematic way. Making the case for a causal relationship is not easy: even if one noticed changing patterns in reversals over time, it is rarely possible to tease apart the influence of docket pressure from that of other factors, such as changes in the quality or makeup of the appeals being brought.¹² These obstacles, along with the judges' hypothesis, are detailed in Part I.

Sometimes, however, history offers researchers a gift. Thanks to an unusual recent event, it is possible to study the effects of a sharp and sudden increase in judicial caseload — a surge that occurred only in certain circuits and not in others. Specifically, I follow the decisions of two circuit courts flooded by tens of thousands of appeals from the federal immigration agency, driven by accelerated deportation "stream-

and Empirical Analysis, J. LEGAL ANALYSIS (forthcoming), available at <http://ssrn.com/abstract=1542834>.

¹¹ One notable exception is Professor Paul Carrington's classic article. See Paul D. Carrington, *Crowded Dockets and the Courts of Appeals: The Threat to the Function of Review and the National Law*, 82 HARV. L. REV. 542, 554 (1969) ("Pressure of time may create a tendency to give greater deference to primary decision makers.").

¹² For an account of composition effects in the rising rate of civil appeals, see CAROL KRAFKA ET AL., *STALKING THE INCREASE IN THE RATE OF FEDERAL CIVIL APPEALS* (1995), available at [http://www.fjc.gov/public/pdf.nsf/lookup/rate_of_appeal.pdf/\\$file/rate_of_appeal.pdf](http://www.fjc.gov/public/pdf.nsf/lookup/rate_of_appeal.pdf/$file/rate_of_appeal.pdf). See also Chris Guthrie & Tracey E. George, *The Futility of Appeal: Disciplinary Insights into the "Affirmance Effect" on the United States Courts of Appeals*, 32 FLA. ST. U. L. REV. 357, 361 & fig.1B (2005) (observing the dual, long-term trends of rising caseloads and declining reversal rates, and noting that "[t]hese two phenomena may be related in any number of ways: one may be a cause of the other or both may be the product of other forces").

lining” after the terrorist attacks of September 11, 2001.¹³ Part II tells this story in more depth. According to the federal courts’ administrative data, this flood of appeals from a single federal agency caused the caseload of each of the two circuits to surge by more than forty percent.¹⁴

To avoid confounding cause with effect, I focus on how outcomes changed *not* in these surging agency cases, but instead in a separate category of cases — civil cases arising from the federal district courts.¹⁵ This broad category, civil appeals, has distinct advantages for study: not only is it unrelated to the source of the surge, but it is also diversified, ranging from civil rights to torts to copyright cases, and hence less easily swayed as a class by changes in any one area of law. Based on this line of reasoning, I exclude from my analysis all criminal appeals, as well as habeas and other prisoner cases; these categories are highly sensitive to even a single change in federal law.¹⁶

Furthermore, three features peculiar to this historical event aid in isolating the impact of caseload, making a causal inference more credible. First, the surge was concentrated in the Second and Ninth Circuits, thus enabling a “natural experiment” interpretation of the findings, with the other circuits providing a baseline for comparison.

¹³ This surge continued unabated through the period of study. My analysis ends in late 2005, just before the Second Circuit’s initiation of the “non-argument calendar” as a special track for processing the immigration appeals (as explained in Part III). See Jon O. Newman, *The Second Circuit’s Expedited Adjudication of Asylum Cases: A Case Study of a Judicial Response to an Unprecedented Problem of Caseload Management*, 74 BROOK. L. REV. 429, 432–34 (2009). My study is thus focused on the period when docket crowding had become so salient that the circuit felt compelled to adopt a major change in how it processed cases — giving up a long tradition of oral argument for nearly every case, *id.* at 433 — but had yet to put the new plan to use.

¹⁴ Between 2001 and 2005, the flow of agency appeals in the Second Circuit rose by roughly 2400 cases per year — a number that is nearly fifty percent of the circuit’s total caseload before the surge. (Accordingly, the share of agency appeals on the docket rose by over thirty percent.) In the Ninth Circuit, too, the surge can be seen as adding about fifty percent to the total caseload (and likewise, the share of agency cases grew by more than thirty percent). See ADMIN. OFFICE OF THE U.S. COURTS, JUDICIAL BUSINESS OF THE UNITED STATES COURTS: 2005 ANNUAL REPORT OF THE DIRECTOR 114 tbl.B-3 (2006), available at <http://www.uscourts.gov/Statistics/JudicialBusiness/JudicialBusiness2005.aspx>.

¹⁵ Even though the immigration appeals are not my object of study, I should add that their own story has begun to be richly examined and vividly told. See, e.g., JAYA RAMJI-NOGALES, ANDREW I. SCHOENHOLTZ & PHILIP G. SCHRAG, REFUGEE ROULETTE: DISPARITIES IN ASYLUM ADJUDICATION AND PROPOSALS FOR REFORM (2009).

¹⁶ Indeed, the period of study encompasses the *Apprendi-Booker* revolution in criminal sentencing. See *United States v. Booker*, 543 U.S. 220 (2005); *Apprendi v. New Jersey*, 530 U.S. 466 (2000). And it also follows closely upon the enactment of both the federal habeas statute and the prisoner litigation statute. See Antiterrorism and Effective Death Penalty Act of 1996 (AEDPA), Pub. L. No. 104-132, 110 Stat. 1214 (codified as amended in scattered sections of the U.S.C.); Prison Litigation Reform Act of 1995 (PLRA), Pub. L. No. 104-134, §§ 801–810, 110 Stat. 1321-66 (1996) (codified as amended in scattered sections of 11, 18, 28, and 42 U.S.C.).

Second, reverse causality is unlikely, given that the federal courts did not create this surge. The judiciary had no say in the deportation streamlining, which the Attorney General abruptly ordered in the aftermath of 9/11. To the contrary, the courts were taken by surprise. Third, by jurisdictional statute, these agency decisions are appealed directly to the circuit courts — bypassing the district courts. The surging agency cases thus cannot have altered the nature of the civil cases by crowding them first at the trial level.

There is a fourth fortuity, one that allows some double-checking of the causal story. In the Second Circuit, another workload crisis had led the chief judge to declare a “judicial emergency” only a few years before, when the departures of judges and political gridlock in filling vacancies had left five out of the thirteen seats vacant. The resulting rise in judicial burdens for the shorthanded court might be expected to have caused effects similar to those of the later surge.

Given the simplicity and transparency of this empirical approach, the raw data essentially tell the story. The picture that emerges is one of lightened scrutiny during the surge. I find that, when flooded by the agency cases, the affected circuit courts began to reverse district court rulings less often — in the *civil* cases. In these circuits, it seems, deference increased, tilting the balance of authority toward the district courts. Moreover, a similar effect occurred during the earlier vacancy crisis in the Second Circuit. These changes, shown in Part III and in the Appendix, can be seen whether the examined sample is all civil cases on appeal, or is limited to subsamples chosen to rule out other influences (such as fluctuating burdens on government lawyers).

What are we to make of this evidence of lower reversals during times of higher judicial burdens? For one thing, it calls attention to the potential for “splits” among the circuits in their levels of scrutiny, caused by differences in caseloads.¹⁷ Moreover, if reversals are falling while the formal standards of review remain unchanged, then the data may be revealing a deference drift that is otherwise eluding notice — and possibly “splits” that are occurring silently.

¹⁷ “Circuit drift” is a term Judge Posner applies to a related phenomenon; as he has noted in another context,

the failure to converge may be an illustration of “circuit drift”: the heavy caseloads and large accumulations of precedent in each circuit induce courts of appeals to rely on their own “circuit law,” as if each circuit were a separate jurisdiction rather than all being part of a single national judiciary enforcing a uniform body of federal law.

Nightingale Home Healthcare, Inc. v. Anodyne Therapy, LLC, 626 F.3d 958, 962 (7th Cir. 2010) (noting divergence of standards for fees under the Lanham Act).

Silence poses special problems for oversight. To the extent the circuits are splitting silently, the usual means for monitoring the work of the courts might be of less use: the Supreme Court's certiorari process, for instance, would not easily detect divergence among the circuits in their de facto levels of deference, so long as they continued to adhere de jure to the same formal standards. And yet the very absence of articulation by the circuit courts may also mean that such drifts or "splits" are unreasoned, thus warranting closer review.

Parts IV and V detail and reimagine potential solutions, focusing on proposals aimed at assigning resources to meet judicial burdens evenly among the circuits, or at allowing judicial resources to flow to the areas of greatest need. Designing more finely tailored solutions, however, demands a much closer look at how cases are handled by the circuit courts than this type of study offers. Notably, this study's unit of analysis is not individual judges, but rather circuit courts as a whole. (Judges are nowhere identified in these data.) What is observed here is the output of a large institution — staff, law clerks, judges, and other officers bound together by norms, culture, habits of workflow, and internal directives. This initial foray thus leaves for future research the hard work of sorting among possible mechanisms for the changes in reversal patterns reported here.¹⁸

Finally, by way of conclusion, I notice a dilemma of judicial integrity: if resource imbalances across circuits are allowed to persist, then attempts to force uniformity by leveling-up reversals in the overtaxed circuits may well come at a cost to another cherished judicial value — the pursuit of correct results.

I. DEFERENCE ADRIFT?

[D]efining the proper scope of review of trial court determinations requires considering in each situation the benefits of closer appellate scrutiny as compared to those of greater deference. — Judge Henry Friendly¹⁹

The bigger the dockets, the less time we spend on the difficult cases and the more mistakes we make. — Judge Harry Edwards²⁰

In the federal judiciary, the courts of appeals are the "courts of the last resort for all but the handful of cases that the Supreme Court will

¹⁸ It leaves open the questions, for instance, of whether the mechanisms include changes in the use of law clerks or staff or visiting judges; time-shifting of when easier or harder cases are decided; changes in how (and how much) case information reaches judicial officers or staff; or other shifts in the court's internal dynamics or case processing. See *infra* Part V, pp. 1145–46.

¹⁹ Henry J. Friendly, *Indiscretion About Discretion*, 31 EMORY L.J. 747, 756 (1982).

²⁰ Harry T. Edwards, *The Role of a Judge in Modern Society: Some Reflections on Current Practice in Federal Appellate Adjudication*, 32 CLEV. ST. L. REV. 385, 403 (1983–1984).

agree to hear.”²¹ Disposing of roughly 60,000 appeals per year,²² these circuit courts have the final say in nearly all of them. And given that their 5000 published opinions each year carry the force of precedent in their jurisdictions,²³ these courts are also among our nation’s most prolific authorities on federal law.

The courts’ workloads are immense, and the judges are keenly aware that time and resources are scarce, making tradeoffs necessary.²⁴ Some judges have described their work as requiring “triage,”²⁵ by which they mean that customary procedures of common law judging — for instance, hearing oral arguments or publishing opinions — must now be limited to a select group of cases.²⁶ One of the constraints creating these tradeoffs is that the federal appeals courts must formally act on all appeals filed, even if the action is as minimal as dismissing a case for being untimely filed.²⁷

Set against the courts’ mandatory jurisdiction over an ever-growing caseload is another hard constraint: frozen judicial capacity. Congress

²¹ Carolyn Dineen King, Lecture, *Challenges to Judicial Independence and the Rule of Law: A Perspective from the Circuit Courts*, 90 MARQ. L. REV. 765, 785 (2007). The “handful” of cases reviewed by the Supreme Court is now about eighty per year. See, e.g., *The Supreme Court, 2009 Term — The Statistics*, 124 HARV. L. REV. 411, 411 tbl.1(A) (2010) (noting that the Supreme Court issued eighty-seven decisions with full opinions in the October 2009 Term); *The Supreme Court, 2008 Term — The Statistics*, 123 HARV. L. REV. 382, 382 tbl.1(A) (2009) (finding seventy-eight decisions with full opinions in the October 2008 Term). For a classic account of caseload concerns at the Supreme Court when its caseload was double the current norm, see Peter L. Strauss, *One Hundred Fifty Cases Per Year: Some Implications of the Supreme Court’s Limited Resources for Judicial Review of Agency Action*, 87 COLUM. L. REV. 1093 (1987).

²² ADMIN. OFFICE OF THE U.S. COURTS, JUDICIAL BUSINESS OF THE UNITED STATES COURTS: 2009 ANNUAL REPORT OF THE DIRECTOR 81 tbl.B (2010), available at <http://www.uscourts.gov/uscourts/Statistics/JudicialBusiness/2009/JudicialBusinesspdfversion.pdf>.

²³ *Id.* at 42 tbl.S-3.

²⁴ Judge Ruggero Aldisert of the Third Circuit described one such tradeoff: “Constraints of time demand the tradeoff. I would rather have adequate time for a decision conference, allowing for the discussion of complex and difficult issues . . . than be forced to shortchange those cases by the process of automatically granting oral argument in every case.” Ruggero J. Aldisert, *Appellate Justice*, 11 U. MICH. J.L. REFORM 317, 321 (1978).

²⁵ “Rational triage” is how Judge Frank Coffin of the First Circuit described the “time-conscious evaluation of cases” for potential publication. FRANK M. COFFIN, ON APPEAL: COURTS, LAWYERING, AND JUDGING 175–76 (1994); see also Lauren K. Robel, *Caseload and Judging: Judicial Adaptations to Caseload*, 1990 BYU L. REV. 3, 9.

²⁶ Time reallocation is necessary if the courts do not permit enough delay or backlog to maintain prior levels of judicial time per case. Circuit courts use the “power of shame” to control delay — “at meetings of the judges, each judge is required to explain the status of every one of the opinions assigned to him that has not been issued within a specified period,” for example ninety days. POSNER, *supra* note 7, at 223.

²⁷ In this sense, they are commonly said to be courts of “mandatory” jurisdiction or review, as opposed to “discretionary” review of the sort exercised by the Supreme Court. Formally discretionary appeals occasionally do come before these courts, but even for such filings, the courts must at least rule on whether the appeal will proceed. See, e.g., 28 U.S.C. § 1292 (2006) (granting federal courts of appeals discretionary jurisdiction over interlocutory appeals).

has declined to create any new appellate judgeships since 1990.²⁸ In that time, the number of appeals has increased by more than 20,000 — or nearly fifty percent.²⁹ Remarking on the glacial growth in congressionally authorized judgeships, Judge Edwards pointed out that “[t]he bankruptcy of supply expansion suggests the need for some sort of rationing of federal judicial time, an undesirable de facto version of which may already be occurring in the courts”³⁰ — and this was a quarter century ago. Tradeoffs, and the need for triage, can have become only more acute in recent years.

A. *The Judges’ Hypothesis*

In light of these tight (and tightening) constraints on the circuit courts, the hypothesis advanced by Judges Gibbons and Posner seems plausible. One might imagine many ways for greater workloads to result in greater deference. Even mechanical changes in how cases are processed — how they are guided through administrative screens, how much oral argument is allowed (if any), whether staff attorneys or law clerks are assigned to them, whether the deciding panel has any visiting judges, and so forth — might affect how close they come to a reversal.³¹ Anywhere along the way, moreover, a rise in time pressure might make time-saving options more attractive. This may even be true at the merits stage, as Judge Ruggero Aldisert has warned, because affirming saves the longer time it takes to reverse.³² And in some cases a decision can be made while deciding fewer issues, as Judge Patricia Wald has explained:

I can tell you that the number of cases that go down on waiver or failure to raise the right point in the right way before the agency or trial court is

²⁸ See *Additional Authorized Judgeships*, ADMIN. OFFICE OF THE U.S. COURTS, <http://www.uscourts.gov/uscourts/JudgesJudgeships/docs/authAppealsJudgeships.pdf> (last visited Jan. 30, 2011).

²⁹ In 2008, 61,104 appeals were filed in the circuit courts (not counting the specialized Federal Circuit). In 1990, there were 40,893. ADMIN. OFFICE OF THE U.S. COURTS, JUDICIAL FACTS AND FIGURES 2008, at tbl.2.2 (2009), available at <http://www.uscourts.gov/uscourts/Statistics/JudicialFactsAndFigures/2008/alljudicialfactsfigures.pdf>.

³⁰ Harry T. Edwards, *The Rising Work Load and Perceived “Bureaucracy” of the Federal Courts: A Causation-Based Approach to the Search for Appropriate Remedies*, 68 IOWA L. REV. 871, 922 (1983).

³¹ Examples of such caseload-driven adaptations are thoughtfully presented in a recent first-hand account by the Ninth Circuit’s then-Clerk of Court. See Cathy Catterson, *Changes in Appellate Caseload and Its Processing*, 48 ARIZ. L. REV. 287 (2006).

³² Ruggero J. Aldisert, *Then and Now — Danger in the Courts*, FED. LAW., Jan. 1997, at 41, 43 (“The danger is that some cases are affirmed rather than reversed because a reversal will require a time-consuming, researched opinion.”). Similarly, Judge Friendly has noted of the district courts, “[I]n these days of crowded dockets there is an inevitable risk of some degree of subconscious bias when [the] decision whether to dismiss a case . . . is made by the judge who will have to [hear] it” Friendly, *supra* note 19, at 754.

too high. In an ideal system of justice, that might not be true, but realistically, time and docket pressures very definitely constrict the judge.³³

Notably, the forfeiture of arguments tends to favor appellees (and affirmances), as it is the appellant who needs to have preserved a supposed error by objecting at the right time.

As is apparent, the shift toward more deferential outcomes described by the judges' hypothesis need not involve any conscious choice by judicial officers or staff to think about a given case differently than before. Yet, as choices go, giving more deference might well be a sensible one. Its rationale would be familiar, with a fine pedigree. Judge Friendly urged "considering in each situation the benefits of closer appellate scrutiny as compared to those of greater deference"³⁴ — and those relative benefits likely depend on the value of time. As Judge Calvert Magruder once put it, "[T]he main reason we on appeal may have a better chance of being right [than do trial judges] is that we have more time for reflection and study."³⁵ Having less time, then, means less advantage; or, in the words of Judge Edwards, "The bigger the dockets, the less time we spend on the difficult cases and the more mistakes we make."³⁶ Moreover, if one recalls that the "circumstances in which Congress or [the Supreme] Court has articulated a standard of deference for appellate review of district-court determinations reflect an accommodation of the respective institutional advantages of trial and appellate courts,"³⁷ then adapting deference to the "conditions of judging" might be a sensible choice.³⁸

B. In Search of Evidence

The judges' hypothesis is plausible, perhaps — but absent evidence, it is also easy to deny. Indeed, one of the judges who proposed the hypothesis has since provided a reason to be skeptical of it. A decade after Judge Posner suggested that "an undesired by-product of the growth in the caseload" might be that "fewer errors made by district

³³ Patricia M. Wald, *Thoughts on Decisionmaking*, 87 W. VA. L. REV. 1, 10 (1984).

³⁴ Friendly, *supra* note 19, at 756.

³⁵ Calvert Magruder, *The Trials and Tribulations of an Intermediate Appellate Court*, 44 CORNELL L.Q. 1, 3 (1958) ("As to the trial judges, we must always bear in mind that they may be as good lawyers as we are, or better. . . . [T]he main reason we on appeal may have a better chance of being right is that we have more time for reflection and study."); *see also* Friendly, *supra* note 19, at 757–58.

³⁶ Edwards, *supra* note 20, at 403.

³⁷ *Salve Regina Coll. v. Russell*, 499 U.S. 225, 233 (1991).

³⁸ I borrow this evocative phrase from Judges Coffin and Katzmman. *See* Frank M. Coffin & Robert A. Katzmman, *Steps Towards Optimal Judicial Workways: Perspectives from the Federal Bench*, 59 N.Y.U. ANN. SURV. AM. L. 377, 377 (2003).

courts are being corrected” by the circuit courts,³⁹ he seems to have reassessed.

“[T]he federal courts of appeals appear to have accommodated the steep increase in caseload per judge relatively painlessly,” he observed just a few years ago, thanks to such “economy measures” as “curtailment of the frequency and length of oral argument,” the “reduced number of cases decided by a published opinion,” and changes in technology and personnel.⁴⁰ As for whether these adaptations have “reduced the quality of the federal judicial output,” he noted that “[n]o general answer is possible,” but “I certainly have no impression that quality has fallen.”⁴¹

What do we make of Judge Posner’s seeming change of heart?⁴² To some, the “relatively painless” view might ring more intuitively true than the deference hypothesis does. Because circuit courts have wide leeway in using time-saving procedures, on many administrative margins, it is not obvious that reversal rates should be yet a further margin showing more than minimal “give.” There are good reasons, after all, to expect that case outcomes would be the last thing to change.

What is more, the data that inspired the original hypothesis go only so far in supporting it. The long-term trends in reversal rates and caseloads cited by Judge Gibbons⁴³ are merely suggestive, at best. The problem is that, as caseloads have grown over time, the composition of appeals is also likely to have changed. Reversal rates will naturally fall as caseloads rise, for instance, if certain classes of weaker appeals are growing at faster rates than is the rest of the docket (consider, as an example, the dramatic rise in the rate of appeal for lawsuits filed by prisoners⁴⁴). Moreover, new federal laws bring new cases to the federal dockets, jointly affecting both caseloads and reversal rates.

For such reasons, studying decades-long, gradual trends in caseloads and reversals is not a fully convincing way to demonstrate the impact of judicial workload on appellate deference. Likewise, simple correlations of reversal rates and caseloads across circuits, without

³⁹ POSNER, *supra* note 7, at 176.

⁴⁰ Richard A. Posner, *Demand and Supply Trends in Federal and State Courts over the Last Half Century*, 8 J. APP. PRAC. & PROCESS 133, 135 (2006).

⁴¹ *Id.*

⁴² Or is it? As to reversals, Judge Posner has more recently noted that “[t]he [Seventh] Circuit doesn’t have one of the heaviest workloads,” adding with a laugh, “Maybe that’s why we reverse so many of the appeals.” Abdon M. Pallasch, *Political Refugees Better Off Right Here*, CHI. SUN-TIMES, Apr. 27, 2009, at 14 (quoting Judge Posner) (internal quotation marks omitted). All joking aside, there may actually be less tension than one might think between falling reversals (in overloaded appeals courts) and the aim of sustaining “the quality of federal judicial output” (by the judiciary as a whole). I explore this idea — intimated by the epigraph from Judge Edwards — in more depth in Part V.

⁴³ See sources cited *supra* note 8.

⁴⁴ See, e.g., KRAFKA ET AL., *supra* note 12, at 9–10.

more, may confound the effects of case composition and workload. These familiar problems of inference can be minimized, however, and next I introduce an empirical approach designed to do so.

II. A NATURAL EXPERIMENT: “THE SURGE”

Overnight we had 14,000 more cases. If we had nothing but immigration cases, we'd be busy morning 'til night. — Judge Barry Silverman⁴⁵

In September 2005, one needed only to walk through the Second Circuit's case management offices to get a feel for the magnitude of this surge: Mountains of briefs had formed in almost every available space. Narrow paths snaked through the valleys, leading to desks fortified on all sides by thick walls of administrative records. — Staff Attorney John Palmer⁴⁶

Imagine a circuit court, overworked and understaffed, encountering a further forty percent increase in its caseload. This surge of cases does not subside. What will the court do? Clearly, something must give. Will it respond to this pressure, as Judges Gibbons and Posner have surmised, by adopting “a posture of increased deference to the rulings of the courts we're supposed to be supervising”⁴⁷ — affirming more often and reversing less?

This question is the starting point for the study's empirical design. A surge of this magnitude has in fact occurred, and I will compare what actually happened in the two courts flooded with extra cases (the Second and Ninth Circuits) against what likely would have happened in those courts absent the surge. Because the counterfactual is unknown, we must do our best to approximate it. The most natural baseline to use is the pattern of outcomes in the same circuit, before the surge.

An improvement is possible if a comparison group happens to be available. In this study, it is. To imagine how outcomes in the surge circuits might have evolved (but for the surge), one can look to the evolution of outcomes in the other, nonflooded circuit courts. This is, in essence, a “natural experiment” design, with the flooded circuits viewed as the experimental group and the remaining circuits viewed as the control group.

⁴⁵ Jill Redhage, *Does This Man Look Like He's Funny? ASU-Educated Judge Uses Humor to Lighten Load of 9th U.S. Circuit Court of Appeals*, TRIBUNE (Mesa, Ariz.), Aug. 31, 2007. Judge Silverman also noted, “This Court of Appeals is a full-time job and then some.” *Id.*

⁴⁶ John R.B. Palmer, *The Nature and Causes of the Immigration Surge in the Federal Courts of Appeals: A Preliminary Analysis*, 51 N.Y.L. SCH. L. REV. 13, 14 (2006–2007). This description is the account of an associate supervisory staff attorney in the Second Circuit during the surge.

⁴⁷ Gibbons, *Maintaining Effective Procedures*, *supra* note 8, at 23; *see also* POSNER, *supra* note 7, at 345.

There is one serious potential concern, of course: if the quality of cases on appeal changed for the same reason as the caseload increase, it would be difficult to isolate the impact of caseload. This is one reason the judges' hypothesis has been hard to verify; whenever caseloads change, it is likely that the quality of those very cases has changed as well. What would be observed is a correlation between caseload and outcomes, but not a causal connection.

In this difficulty lies an advantage of studying the surge: I am able to focus on a source of docket pressure (direct appeals from a single federal agency's decisions) that is otherwise unrelated to the cases whose outcomes I measure (civil cases being appealed from the district courts). This separation supports the assumption that the only way for the first category of cases to influence the other is through docket crowding at the court of appeals. To see this more clearly, it is useful to understand first the peculiarities of the surge and its origins.

A. *The Unusual Origins of the Surge*

Shortly after 9/11, in February 2002, Attorney General John Ashcroft pledged at a news conference that the Department of Justice (DOJ) would quickly clear out a deportation backlog, consisting of some 56,000 foreign nationals awaiting hearings before the DOJ's Board of Immigration Appeals (BIA).⁴⁸ In March, the BIA chairperson extended a special "streamlined" review process to all asylum and deportation appeals (that is, most of the BIA's cases).⁴⁹

As the federal judiciary newsletter observed, "Almost immediately, BIA doubled production, sending a deluge of petitions for review into the U.S. courts of appeals . . ."⁵⁰ Moreover, the appeals rate from BIA decisions soared — because more of these decisions upheld depor-

⁴⁸ Press Release, Dep't of Justice, Department of Justice Unveils Administrative Rule Change to Board of Immigration Appeals in Order to Eliminate Massive Backlog of More than 56,000 Cases (Feb. 6, 2002), available at http://www.justice.gov/opa/pr/2002/February/02_ag_063.htm.

⁴⁹ See Memorandum from Lori L. Scialabba, Acting Chairman of the BIA, to all BIA Members (Mar. 15, 2002), in DORSEY & WHITNEY LLP, STUDY CONDUCTED FOR AMERICAN BAR ASSOCIATION COMMISSION ON IMMIGRATION POLICY, PRACTICE AND PRO BONO RE: BOARD OF IMMIGRATION APPEALS: PROCEDURAL REFORMS TO IMPROVE CASE MANAGEMENT app. at 22 (2003). The streamlined process authorized "affirmance[s] without opinion" by any single member of the appeals board; it also made single-member decisions the norm (rather than three-member decisions, as had previously been the standard). *Id.*; see also Palmer, *supra* note 46, at 19.

⁵⁰ *Immigration Appeals Surge in Courts*, THIRD BRANCH (Admin. Office of the U.S. Courts, Washington, D.C.), Sept. 2003, at 5, 6.

tation, and over half were unexplained summary orders⁵¹ — thereby sustaining the surge.⁵²

How this continuing flood of cases affected the dockets of the Second, Ninth, and comparison circuits is shown in Figures 1, 2, and 3. The comparison group consists of all other circuits from First to Eleventh; that is, all federal appeals courts are included in the analysis except the D.C. Circuit and the Federal Circuit, both of which have specialized dockets.⁵³ (In these graphs, each dot represents the number of filings of appeals in a given quarter year. The lower set of dots shows quarterly filings in BIA cases, and the upper set shows filings in all other cases combined — civil, criminal, habeas, and so forth.) As these figures vividly show, the surge was far more severe in the Second and Ninth Circuits than in the other courts.

This surge continued unabated for several years, and the docket pressure it caused led to adaptations in the courts' procedures. Most notably, in October 2005, the Second Circuit adopted a "non-argument calendar" (NAC) to expedite BIA appeals, expressly in the interest of docket relief. In doing so, it abandoned a long-held tradition of allowing the option of oral argument in nearly all cases.⁵⁴

B. *Toward a Causal Story*

Several key features of this surge help in isolating the causal effect of caseload on outcomes. First, the surge of agency cases was concentrated in two regional appeals courts, the Second and Ninth Circuits. Other circuits were much less affected (see Figures 1, 2, and 3). The Second Circuit and the Ninth Circuit contain the locations where roughly three-quarters of the foreign nationals whose cases constituted the surge were initially processed by an immigration judge,⁵⁵ making those two circuits the statutory venues for their appeals.⁵⁶ This pecu-

⁵¹ DORSEY & WHITNEY LLP, BOARD OF IMMIGRATION APPEALS: PROCEDURAL REFORMS TO IMPROVE CASE MANAGEMENT, SUMMARY OF FINDINGS AND CONCLUSIONS 2 (2003), available at http://www.dorsey.com/files/upload/Summary-Conclusion_DorseyABAStudy.pdf.

⁵² See Palmer, *supra* note 46, at 32 fig.3.

⁵³ The Federal Circuit is not included in the data from the Administrative Office of the U.S. Courts (AO). Including the D.C. Circuit in the analysis makes no substantive difference in the results, and given the circuit's highly unusual docket, excluding it likely allows a more credible presentation for most readers.

⁵⁴ The Second Circuit's tradition of allowing the option of oral argument did not extend to appeals from incarcerated persons. See Newman, *supra* note 13, at 432–34.

⁵⁵ The Second Circuit is dominated by New York and also covers Connecticut and Vermont. The Ninth Circuit covers the Pacific states and their neighbors: California, Oregon, and Washington; Alaska and Hawaii; and Arizona, Nevada, Idaho, and Montana. It also covers Guam and the Northern Mariana Islands. See *Court Locator*, ADMIN. OFFICE OF THE U.S. COURTS, http://www.uscourts.gov/court_locator.aspx (last visited Jan. 30, 2011).

⁵⁶ See 8 U.S.C. § 1252(b)(2) (2006).

liar pattern sets up the “natural experiment” — with two circuits receiving the (metaphorical) experimental treatment and the remaining circuits serving as the placebo or control group. In econometric terms, it enables a differences-in-differences analysis.

FIGURE 1: APPEALS FILED IN THE SECOND CIRCUIT 1994–2005

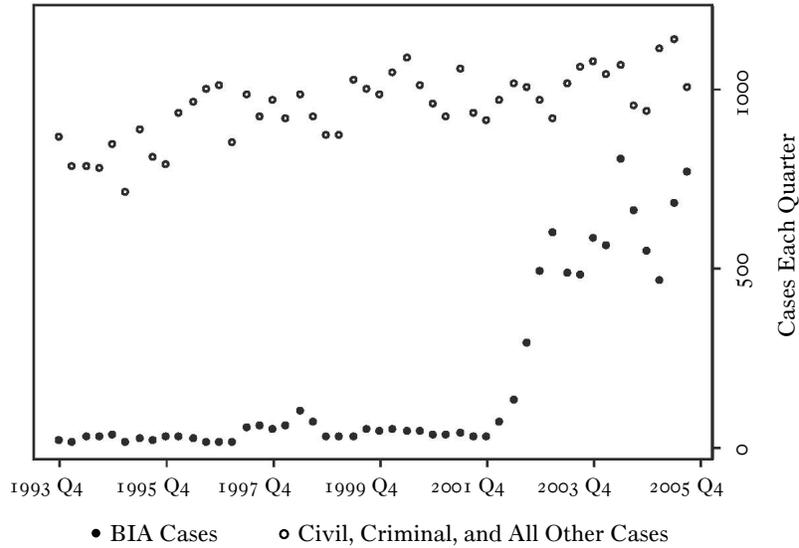


FIGURE 2: APPEALS FILED IN THE NINTH CIRCUIT 1994–2005

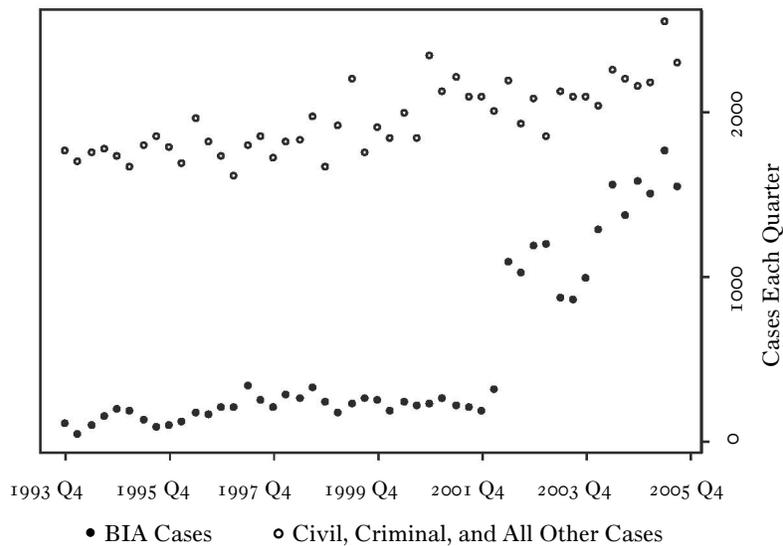
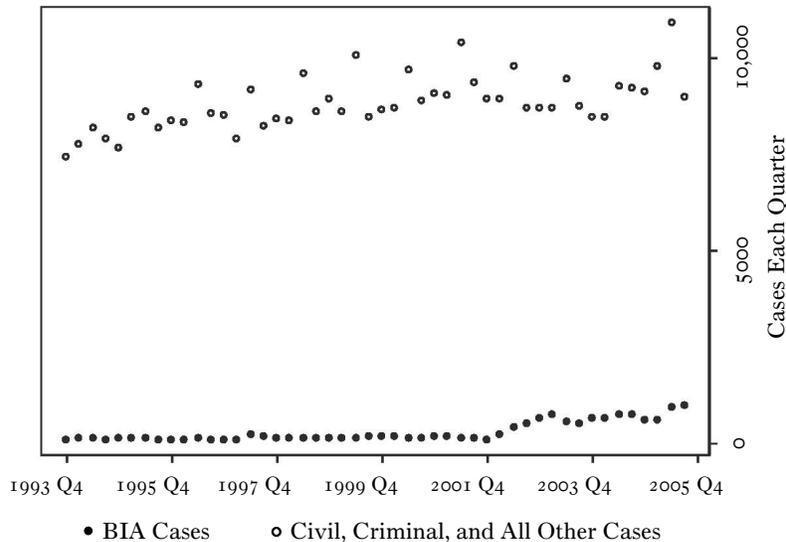


FIGURE 3: APPEALS FILED IN THE COMPARISON CIRCUITS
1994–2005



Moreover, the federal courts were not involved in the immigration agency’s decision to clear its deportation backlog. This fact relieves some potential concerns about reverse causality.⁵⁷ BIA officers are civil servants at DOJ, not a part of the judiciary; and by all indications “streamlining” was an internal agency decision.⁵⁸ Judges expressed surprise; for instance, Ninth Circuit Judge Dorothy Nelson reported: “It’s just extraordinary. I’ve been on the court for 25 years, but I’ve never seen a rush . . . overwhelming us like this.”⁵⁹

Third, these agency appeals completely bypass the federal district courts. Instead, in accordance with a jurisdictional statute, they are appealed directly from the federal agency to the circuit courts.⁶⁰ This fact lessens the concern that the surge might have changed the composition or quality of cases being appealed from the district courts, by crowding them during that earlier stage as well. (To emphasize: I exclude the surging agency appeals themselves from all my outcome

⁵⁷ Also, none of the immigration agency’s decisionmakers seemed to recognize that streamlining might cause a flood of cases in the federal courts. See DORSEY & WHITNEY LLP, *supra* note 49, at 39–41.

⁵⁸ *Id.* at 16–19.

⁵⁹ Solomon Moore & Ann M. Simmons, *Immigrant Pleas Crushing Federal Appellate Courts*, L.A. TIMES, May 2, 2005, at A1 (alteration in original).

⁶⁰ 8 U.S.C. § 1252(b)(2).

measures; they are not classified as “civil” cases in the data, and more to the point, their overall quality is plainly not comparable before and during the surge.)

C. A Second Experiment

As it happens, another “natural experiment” is available in one of our studied circuits to serve as a sort of double check on the causal story being told about the surge. In the spring of 1998, the Second Circuit’s chief judge was forced to declare a “judicial emergency” when five out of thirteen judicial positions on the circuit had become vacant — and were left unfilled by a Senate hostile to the President’s nominees. At that point, the circuit’s vacancy rate had risen to nearly forty percent. This crisis had begun to build in 1996; four vacancies would open by the following summer.⁶¹ By the fall of 1997, a fifth vacancy was anticipated,⁶² prompting the chief judge to warn in testimony before a congressional committee that a declaration of emergency measures was imminent.⁶³ The declaration came in early 1998,⁶⁴ along with the fifth vacancy. Then the situation ended suddenly in late 1998, when a Senate breakthrough allowed four of five vacancies to be filled by long-stalled nominees.⁶⁵

Beyond offering a second historical shock, this earlier event in the Second Circuit also presents a few advantages: Because there was both

⁶¹ Three judges retired from regular active service, taking senior status: Judge Altamari in January 1996, Judge Miner in January 1997, and Judge Newman in July 1997. The fourth, Judge Mahoney, died in October 1996. See *Biographical Directory of Federal Judges*, FED. JUD. CENTER, <http://www.fjc.gov/public/home.nsf/hisj> (last visited Jan. 30, 2011).

⁶² Judge McLaughlin took senior status in March 1998. *Id.*

⁶³ *Considering the Appropriate Allocation of Judgeships in the U.S. Courts of Appeals for the Second and Eighth Circuits and the First, Third, and Federal Circuits: Hearing Before the Subcomm. on Admin. Oversight and the Courts of the S. Comm. on the Judiciary*, 105th Cong. 47 (1997) (statement of Hon. Ralph K. Winter, C.J., U.S. Court of Appeals for the Second Circuit) (warning he may need “to certify an emergency under 28 U.S.C. § 46(b)”). This vacancy threshold seems consistent with what Judge Niemeyer, a Fourth Circuit judge, has recently said about vacancies in his court (which has fifteen authorized seats): The circuit “can operate with twelve or thirteen judges without a problem. But, if we are reduced to ten judges, things will get a little dicey, and it will be difficult to operate efficiently.” Paul Mark Sandler, *A Conversation with Judge Niemeyer*, in *APPELLATE PRACTICE FOR THE MARYLAND LAWYER: STATE AND FEDERAL § II-7* (Paul Mark Sandler & Andrew D. Levy eds., 3d ed. 2007) (quoting Judge Niemeyer), available at APML MD-CLE 7-73 (Westlaw).

⁶⁴ See *Wenger v. Canastota Cent. Sch. Dist.*, 146 F.3d 123, 124 n.*** (2d Cir. 1998) (recognizing the existence of a judicial emergency).

⁶⁵ Judges Straub, Pooler, and Sack began service in June 1998, and then-Judge Sotomayor started in October 1998. See *Biographical Directory of Federal Judges*, *supra* note 61. Senator Patrick Leahy has suggested that the extra delay for Judge Sotomayor was intended by his Republican colleagues to prevent her from being nominated that same summer for a possible Supreme Court vacancy. See *Nominations: Hearing Before the S. Comm. on the Judiciary*, 111th Cong. (2010) (statement of Sen. Patrick Leahy, Chairman, S. Comm. on the Judiciary), available at http://judiciary.senate.gov/hearings/testimony.cfm?id=4533&wit_id=2629.

a collapse and a rebound in judicial resources, we can more confidently attribute a causal effect if the outcomes also show a down-and-up “V” pattern. Moreover, because both the opening and the filling of these vacancies shifted the circuit toward a greater share of Democrat-appointed judges,⁶⁶ it would be hard to attribute such a “V” pattern in case outcomes to a simple story about “political” shift on the bench. It cannot be that both the sudden fall in reversals and the sudden rebound are attributable to shifts toward a more Democrat-appointed court.

Finally, this extra episode may also inform whether drifting in silence is a likely possibility: if a sharp “V” pattern is seen in reversal rates, it is implausible that formal doctrinal changes in standards of review are the reason; such a story would require dramatic changes in doctrine in one direction, and then equally severe changes in the other direction, within a span of two years.

III. FINDINGS: LIGHTENED SCRUTINY

The signs of lightened scrutiny are easily seen in the raw data. I therefore present these findings mainly in graphical form. In the Appendix, I also illustrate how econometric methods can be used to confirm and measure what our eyes are seeing here, to control for additional background factors, and to focus on specific subsamples.

In this Part, I first describe the data. Next, I show the change in patterns of reversals from the Ninth Circuit, which presents a clearer before-and-after story. I then turn to the Second Circuit, which has a more complicated history during the period of study — but complicated in a good way, offering a second “experiment” (an earlier crisis of judicial burdens) that can be used as a double check for interpretations of what is seen in the Second and Ninth Circuits later during the surge.

A. The Data

The data sample for both the graphs and the regressions consists of civil appeals arising from the district courts, as documented by the Administrative Office of the U.S. Courts (AO).⁶⁷ How I have limited this sample and cleaned the data is detailed in the Data Appendix.

⁶⁶ Of the original five judges, all but Judge Newman were appointed by Republican Presidents. See *Biographical Directory of Federal Judges*, *supra* note 61. And the 1998 appointees were nominated by President Clinton.

⁶⁷ General background on the AO data is detailed with care in Theodore Eisenberg & Margo Schlanger, *The Reliability of the Administrative Office of the U.S. Courts Database: An Initial Empirical Analysis*, 78 NOTRE DAME L. REV. 1455 (2003). Their informative essay focuses on certain limitations of the district court data (not relevant to the appeals data used here) and also catalogs the growing use of AO data in academic literature.

Most notably, habeas and other prisoner suits are excluded: like criminal cases, such suits are highly vulnerable to shocks in federal law, complicating the task of isolating caseload effects; and here, the period of study begins shortly after passage of statutes concerning habeas and prisoner litigation — namely, the Antiterrorism and Effective Death Penalty Act of 1996⁶⁸ (AEDPA) and the Prison Litigation Reform Act⁶⁹ (PLRA) — and is interrupted by the *Apprendi-Booker* line of criminal sentencing decisions.⁷⁰ The portfolio of civil cases, by contrast, is widely diversified and less responsive to shocks in any single area of law.

The period I analyze begins in late 1997 and ends in late 2005.⁷¹ During this time, the number of authorized circuit judgeships remained constant (again, Congress has not increased the number of circuit judges for twenty years). The federal judiciary's staffing budget also did not appreciably increase. The starting point for this period, the fourth quarter of 1997 (the beginning of the 1998 statistical reporting year), is the earliest date at which this data source indicates whether a case involved a pro se litigant. It will be useful, as noted in the Appendix, to check our findings in a sample limited to cases in which all parties are represented by a lawyer (that is, excluding the pro se cases), seeing as pro se cases receive special administrative treatment in the circuit courts.⁷²

The study period closes in the third quarter of 2005 (at the end of the 2005 statistical reporting year) — well before a highly controversial change in policy occurred throughout the circuit courts: a new federal rule governing the citation of unpublished opinions, which was approved in 2006 and took effect at the start of 2007.⁷³ The observed patterns in outcomes thus cannot be due to that policy change.

Notably, the data period shown also ends just before the Second Circuit began to use its new non-argument calendar (or NAC, which in effect denies oral argument for most asylum and deportation appeals),⁷⁴ which it created expressly to reduce the docket pressure

⁶⁸ Pub. L. No. 104-132, 110 Stat. 1214 (codified as amended in scattered sections of the U.S. Code).

⁶⁹ Pub. L. No. 104-134, §§ 801–810, 110 Stat. 1321–66 (1996) (codified as amended in scattered sections of 11, 18, 28, and 42 U.S.C.).

⁷⁰ These sentencing decisions include *United States v. Booker*, 543 U.S. 220 (2005); and *Apprendi v. New Jersey*, 530 U.S. 466 (2000).

⁷¹ On some graphs I will also show data going farther back. This is especially useful for seeing the effects of the Second Circuit's earlier vacancy crisis, as explained above.

⁷² See JUDITH A. MCKENNA, LAURAL L. HOOPER & MARY CLARK, FED. JUD. CENTER, CASE MANAGEMENT PROCEDURES IN THE FEDERAL COURTS OF APPEALS 52 (2000), available at [http://www.fjc.gov/public/pdf.nsf/lookup/caseman1.pdf/\\$file/caseman1.pdf](http://www.fjc.gov/public/pdf.nsf/lookup/caseman1.pdf/$file/caseman1.pdf).

⁷³ Federal Rule of Appellate Procedure 32.1(a) requires that all circuits permit the citation of any unpublished opinion issued on or after January 1, 2007. FED. R. APP. P. 32.1(a).

⁷⁴ 2D CIR. R. 34.2.

brought on by the surge. The data shown, therefore, do not reflect whatever relief this process may have provided.

A few limitations should be noted. First, for the reasons given above, this study considers only civil cases. It is something of an irony that possibly the most intense dispute over appellate review in recent years — the *Apprendi-Booker* line of cases about review of criminal sentencing — weighs against including criminal cases in this study.⁷⁵ My findings on civil appeals may have little bearing on criminal appeals, especially as many circuits give calendaring priority to cases involving incarcerated parties.⁷⁶

Second, information about case characteristics is scarce in this database. The problem of possible unobserved factors, or omitted-variables bias, is reduced by the “natural experiment” study design — but not eliminated. As seen in the Appendix, econometric methods are available for reducing this concern further (for instance, by separately analyzing subcategories of cases and by including indicators for each circuit and for each time period). Nonetheless, this shortcoming leaves noise in the data and precludes the study of some interesting questions.

Finally, this study has nothing to say about individual judges. For one thing, no judges are identified in the AO data. More importantly, many of a circuit court’s outputs are collaborative products.⁷⁷ None of the findings reported here should be read as attributing changes in outcomes to any given layer of the court’s decisional apparatus, much less to any individuals. (It is good to keep in mind that even “signed” judicial opinions result from the combined efforts of law clerks or staff attorneys and the authoring judge, often with input from other chambers.) Moreover, many mechanisms by which caseload pressure might affect case outcomes can be imagined — evolution in the use of law clerks or staff or visiting judges, shifts in the timing of workflow or in who gets what information, revised voting methods, or other subtle changes in a court’s internal dynamics.⁷⁸ It remains for future work to sort among them.

⁷⁵ Again, this is because the data on reversals in criminal cases are hard to interpret in this period; it cannot be said (with *Apprendi* decided in 2000 and *Booker* in 2005) that the inherent characteristics of federal criminal appeals were steady during the studied period. (Habeas and other prisoner cases suffer a similar problem, given the enactment of the AEDPA and the PLRA near the start of the period of study, and the changing case law concerning those statutes in the years following.)

⁷⁶ See MCKENNA, HOOPER & CLARK, *supra* note 72, at 51.

⁷⁷ See Henry Paul Monaghan, *Taking Bureaucracy Seriously*, 99 HARV. L. REV. 344, 346 (1985) (reviewing RICHARD A. POSNER, *THE FEDERAL COURTS: CRISIS AND REFORM* (1985)).

⁷⁸ I thank Judge Newman for suggesting specific possibilities (without endorsing any of them); he also emphasized that conscious change in judicial behavior is an unlikely one.

B. Revealed Deference

During the surge period, both the Second and Ninth Circuits show marked declines in how often they reversed decisions of the district courts. Figures 4, 5, and 6 show the number of cases resulting in reversals, remands, and partial reversals combined.⁷⁹ I will call this combination “reversals,” for short; it yields a measure of those cases in which the appeals court chose to undo at least one aspect of a trial court ruling.⁸⁰ Figures 4 and 5 present this measure for each of the two flooded circuits, separately.

To consider the Ninth Circuit first: Figure 4 shows a clear drop-off in the Ninth Circuit’s reversals during the surge (to the right of the solid line, which marks the start of the surge). It is easy to see that the number of reversals begins to fall, after a time lag during which the surge grows and the agency cases begin occupying the work time of the court’s staff and officers.

In the Second Circuit, as already noted, the story is more complicated. As in the previous graph, the drop in reversals during the surge is also apparent for this circuit, as seen in Figure 5: reversals are higher in the stretch from late 1998 to early 2003 than afterwards. Between 1996 and 1998, this circuit faced a separate crisis. When judicial vacancies grew to five out of thirteen seats on the active bench, the chief judge declared a “judicial emergency,” specifically citing the problem of having too few judges to fill the panels needed to hear all the cases. Four vacancies were then suddenly filled in the second half of 1998. As explained in section II.C, this period (as marked off by the dashed vertical lines) may be seen as a second “experiment” for assessing the impact of judicial overload. Indeed, the data show a marked decline in reversals during this period — and a sharp rebound once the vacancies were filled.

⁷⁹ By partial reversals, I mean any decision listed as “reversed in part,” including those which were also “affirmed in part.”

⁸⁰ Moreover, this combined measure allows greater comparability, as the circuits tend to vary in their usage of the terms “reversal” and “remand” (at times interchanging them). Judge Newman, formerly Chief Judge of the Second Circuit, has documented these varying usages, see Jon O. Newman, *Decretal Language: Last Words of an Appellate Opinion*, 70 BROOK. L. REV. 727, 729–31 (2005), and he has also used the same shorthand I am using here, see Jon O. Newman, *A Study of Appellate Reversals*, 58 BROOK. L. REV. 629, 629 (1992). The coding of the administrative data does not distinguish between decisions that are “reversed” and those that are “vacated.”

FIGURE 4: REVERSALS FALL IN THE NINTH CIRCUIT DURING THE SURGE

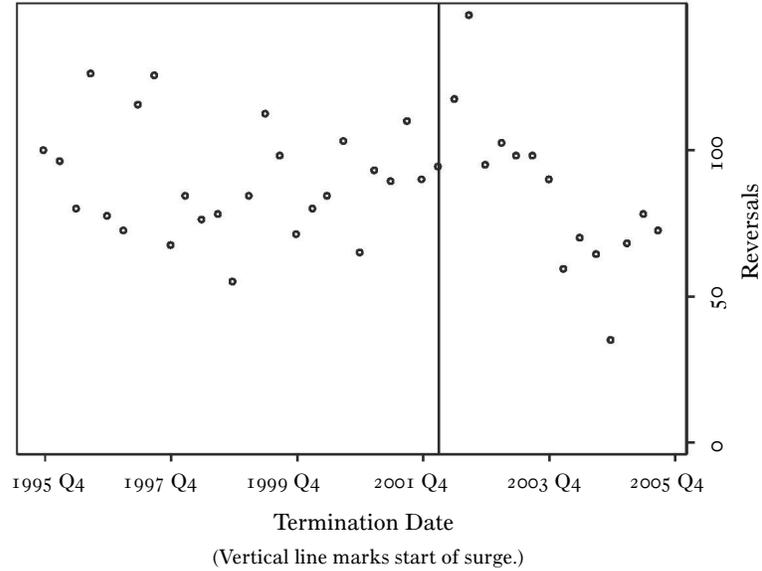


FIGURE 5: REVERSALS FALL IN THE SECOND CIRCUIT DURING BOTH THE CRISIS AND THE SURGE

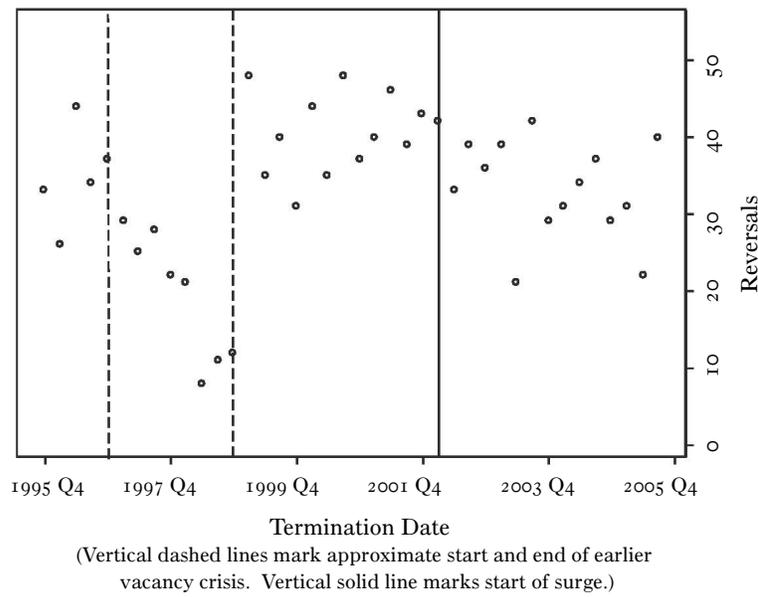
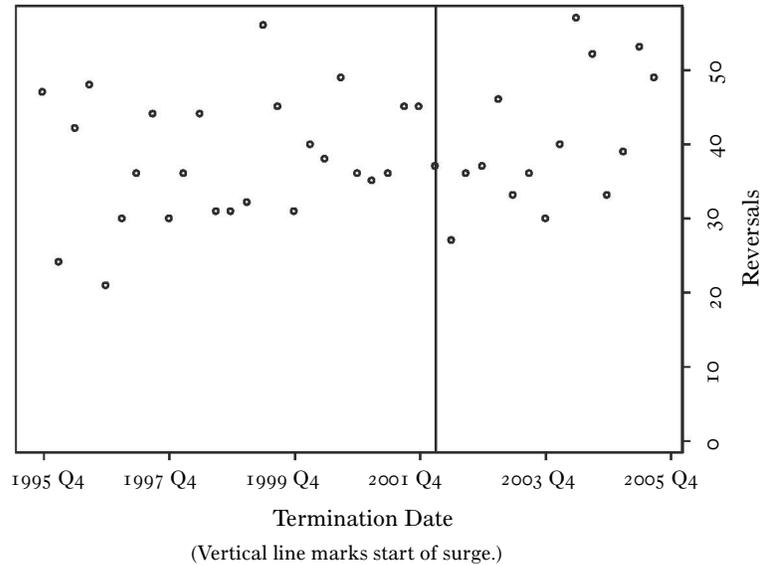


FIGURE 6: REVERSALS REMAIN STEADY IN THE THIRD CIRCUIT



To offer a contrast, Figure 6 shows reversals in the Third Circuit, which borders the Second Circuit and is conveniently well matched for comparison: Each is anchored by a major metropolitan area (New York City in the Second Circuit, and Philadelphia and parts of New Jersey in the Third Circuit). Each covers three states (New York, Connecticut, and Vermont; as compared to Pennsylvania, New Jersey, and Delaware). They cover similar numbers of districts (five, as compared to six). They have nearly the same number of active judgeships (thirteen, as compared to fourteen). And the timing of President George W. Bush's appointments of judges to these circuits was similar within the study period. What does differ between the Second and Third Circuits is the time path of reversals during both the earlier vacancy crisis and the later surge. In contrast to the fluctuations in the Second Circuit, the Third Circuit's reversals hover at roughly the same level throughout the period of study.

This contrast is seen on a larger scale in Figures 7 and 8, which show reversal rates in the Second and Ninth Circuits (taken together) in comparison to all the remaining circuits (taken together). The measure here is not the raw number of reversals, but rather the percentage reversed among all terminations. Each dot represents this percentage in a given quarter — that is, among all dispositions of appeals in a three-month period, what share were reversals? A fall in this measure in the Second and Ninth Circuits is seen both during the surge (as traced out by the fitted curve) and also during the Second Circuit's earlier crisis (to the left of the vertical dashed line).⁸¹ Reversal rates in the comparison circuits, by contrast, seem fairly flat over time.⁸²

Another way to view these differences is from the perspective of an appeal at the time it begins; and Figures 9 and 10 show the outcomes of appeals by their date of *filing* (not termination). That is, among all appeals filed during a given quarter year, what share eventually resulted in reversal? Again, the downturn in reversal rates in the Second and Ninth Circuits for appeals filed during the surge is noticeable, as is an earlier dip corresponding to the Second Circuit's earlier crisis.

C. *Alternative Stories?*

A few words about possible alternative explanations are in order. First, it is worth a reminder that recent, high-profile rule changes cannot account for the results. The new national rule concerning the citing of unpublished opinions did not take effect until 2007, after the period of study. Up until then, both the Second and Ninth Circuits adhered to longstanding (and unchanging) local rules introduced in the 1970s.⁸³ Likewise, the Second Circuit's creation of the non-argument calendar, or NAC, did not occur until October 2005, just after the study period.⁸⁴

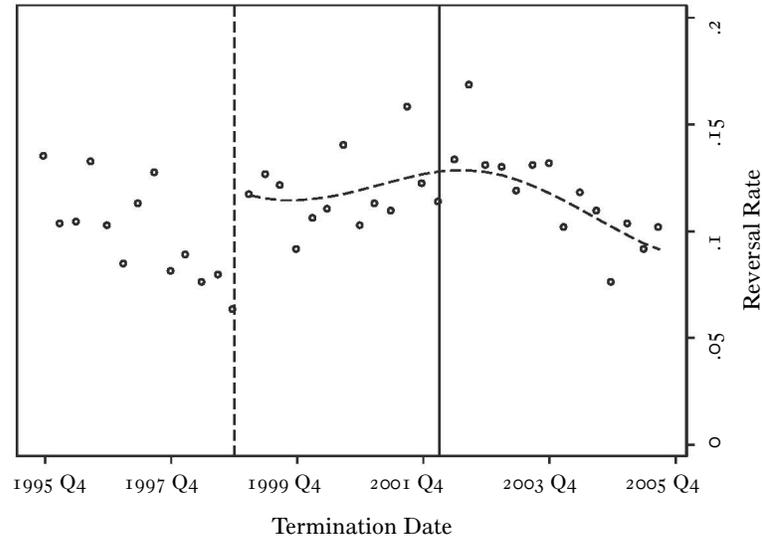
⁸¹ I avoid presenting a fitted line in the earlier crisis period (though the shape may be obvious) so as not to encourage reading too much into how far this measure falls, given that it pools both circuits together and only the Second Circuit could have been affected by the crisis in that circuit. This is less of a worry for the surge period, during which the proportional increases in caseload in these two circuits were very similar; hence the fitted curve there. These are aesthetic, and not substantively important, choices.

⁸² Moreover, no other individual circuit matches the pattern of outcome changes seen in the Second and Ninth Circuits during the surge period.

⁸³ See FED. R. APP. P. 32.1; Robert Timothy Reagan, *Citing Unpublished Federal Appellate Opinions Issued Before 2007*, FED. JUD. CENTER (Mar. 9, 2007), [http://www.fjc.gov/public/pdf.nsf/lookup/citrules.pdf/\\$file/citrules.pdf](http://www.fjc.gov/public/pdf.nsf/lookup/citrules.pdf/$file/citrules.pdf). Of the circuits studied, only the First Circuit changed its citation rules during the period of study; its outcome measures show no noticeable changes coinciding with its rule change. The D.C. Circuit also changed its rules during this period, but it is excluded from this study. (Again, including the D.C. Circuit makes virtually no difference in the results.)

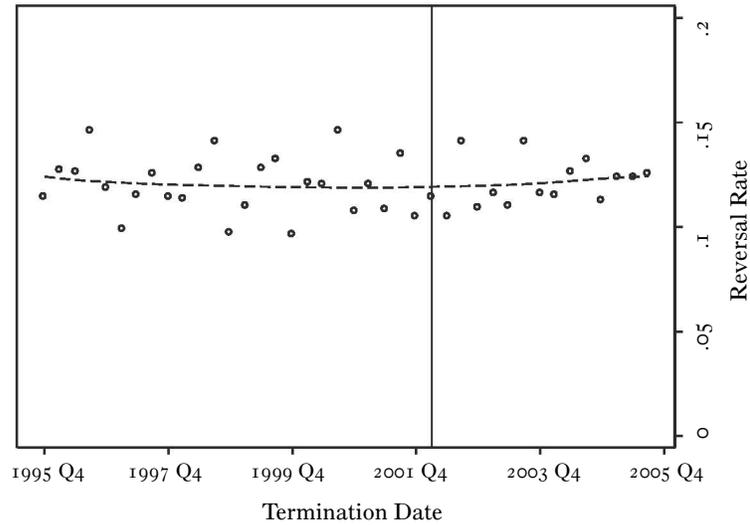
⁸⁴ See 2D CIR. R. 34.2.

FIGURE 7: REVERSAL RATES FALL DURING THE SURGE
IN THE SECOND AND NINTH CIRCUITS



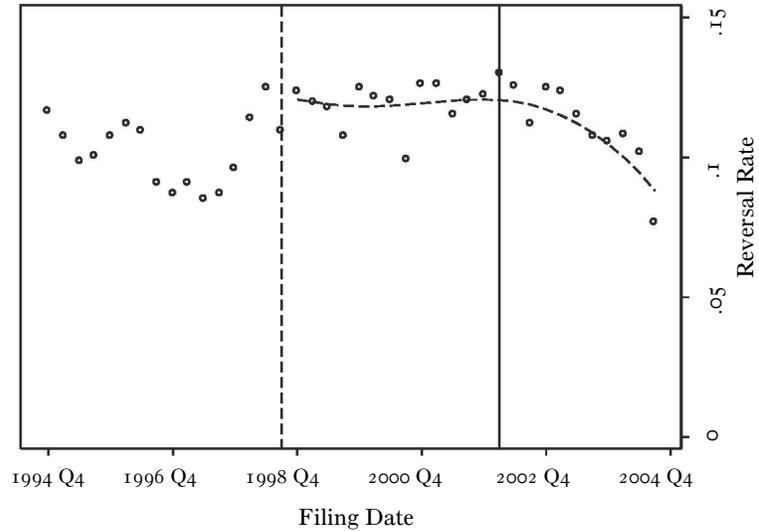
(Vertical dashed line marks approximate end of earlier Second Circuit vacancy crisis. Vertical solid line marks start of surge in filings.)

FIGURE 8: REVERSAL RATES REMAIN STEADY
IN THE COMPARISON CIRCUITS



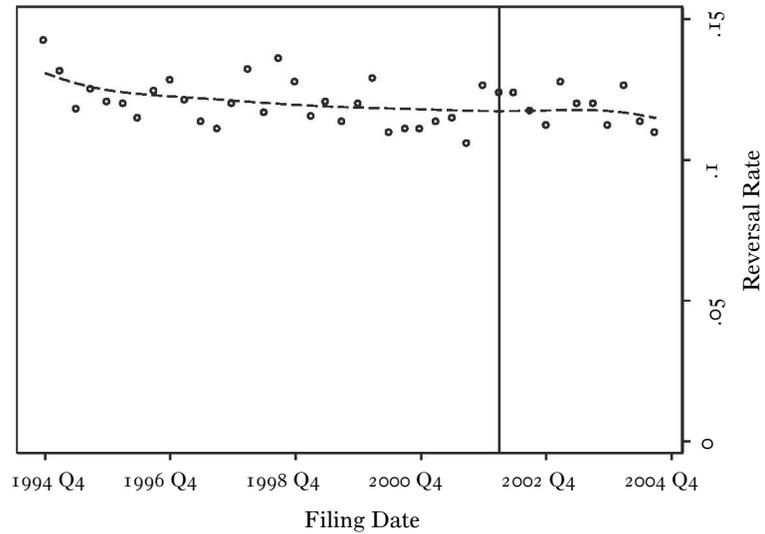
(Vertical line marks start of surge.)

FIGURE 9: REVERSAL RATES ARE LOWER FOR APPEALS FILED DURING THE SURGE IN THE SECOND AND NINTH CIRCUITS



(Vertical dashed line marks approximate end of earlier Second Circuit vacancy crisis. Vertical solid line marks start of surge in filings.)

FIGURE 10: REVERSAL RATES ARE STEADY IN THE COMPARISON CIRCUITS



(Vertical line marks start of surge.)

Second, could the observed divergences have resulted from changes in the composition of civil cases in the Second and Ninth Circuits? This seems unlikely. The pattern of civil appeals filings in the Second and Ninth Circuits shows no abrupt shifts; rather, it matches the gradual pattern of smooth, mild decline seen in the comparison circuits. (To emphasize: habeas and prisoner suits are excluded from my sample of “civil” cases.) Moreover, as seen in the Appendix, the findings persist whether the sample is all civil cases, or limited to cases in which all parties have counsel (that is, excluding pro se cases), or further limited to private cases (excluding cases in which the federal government is a party).

Notably, the smooth path of appeals filings suggests that would-be appellants have not reacted in large numbers to the outcome changes reported here. It is possible that the changes, to the extent they were noticeable to litigants, were not salient enough to affect many parties’ appeals decisions or strategies. A further possibility is that most parties simply did not know. Even if they were inclined to try, calculating up-to-the-moment changes in average reversal rates for a given type of case or on a given issue is probably not easy for most parties.⁸⁵ Along these lines, it may be telling that the Second Circuit’s earlier vacancy crisis, which not only generated sharp changes in reversal rates but also was publicly announced, likewise did not seem to perturb the pattern of filings of civil appeals.

The double check provided by the earlier Second Circuit vacancy crisis also helps to allay a third worry: that the observed changes in reversals are the result of a changing “political” composition of the bench. During its earlier crisis, the Second Circuit twice shifted toward being a more Democrat-appointed bench: both as the vacancy crisis started (four of five departures were Republican appointees) and as it ended (the new appointments were all by President Clinton). Thus the change in “politics” went in the same direction each time. And yet the changes in reversals went in opposite directions, first collapsing, then rebounding. It would be implausible to link these patterns causally. Further evidence against a simple “politics” story is found in the matched comparison between the Second and Third Circuits. Both these circuits gained Bush-appointed judges at roughly the same rate, and yet their patterns of reversals were in marked contrast during those years. A similar logic also applies, writ large: the “natural experiment” (or difference-in-differences) setup already accounts for “political” change, to an extent, because such changes tend to be

⁸⁵ Consider a party deciding whether to file an appeal, say, in 2005. Searching on Lexis or Westlaw might have turned up very few cases addressing the same issue and decided in the surge period (after mid-2002); probably this litigant would have had to look further back, and include cases before the surge, to get a usable sample.

shared by all the circuits, seeing as their new judges are all appointed by the same President at any given time. It is thus unsurprising that adding further controls for the “political” makeup of the district courts and circuit courts in my statistical analysis (in the Appendix) turns out to be unimportant.

Finally, could the trauma of 9/11 itself have created stresses on the work of the Second Circuit, thus accounting for the changes observed there? The pattern of the changes suggests otherwise. A direct effect of 9/11 would likely have caused a sharp change, and then a rebound, in observed outcomes after 2001.⁸⁶ But the observed changes grow, rather than diminish, over time — consistent with increasing pressure from the continuing surge. Moreover, one might expect the stress of 9/11 to have affected the Third Circuit (which includes New Jersey and Pennsylvania), the Fourth Circuit (which includes Virginia), and the D.C. Circuit; but their case outcomes evolved very differently from those of the Second. The Ninth Circuit’s changes are similar to the Second Circuit’s but are much less likely to be due to direct effects of 9/11.

IV. PROBLEMS FOR THEORY AND POLICY

This study surfaces a series of problems for theory and policy, ranging from simple to more subtle. Imagining solutions and answers necessarily engages broader debates — about the role of Congress and the President in staffing the judiciary, about the use of doctrine and precedent in governing judicial review, and even about the competing values that comprise judicial integrity. In the following sections, I will touch on these concerns, but only lightly, and I will not try to survey the literature of reform proposals for the judiciary (a genre unto itself, vast even as to the circuit courts alone). Instead, my primary aim here is to highlight those lines of questioning that are specially motivated by this study.

A brief survey of the concerns raised may help set the scene. First, there is the most immediate matter of lightened scrutiny. The evidence suggests that the studied appeals courts began to call out fewer problems in trial court rulings, in civil cases, when they became overburdened by a wholly separate set of cases. Whatever the precise mechanisms of this change may be, one lesson is easily seen: These data serve as a reminder of how “legal spillovers” can occur in our system of generalist courts. A single federal agency’s change in internal pro-

⁸⁶ Tellingly, the *total* number of civil appeals terminated was lower than usual in third-quarter 2001 (ending September 30) — but it immediately rebounded by the next quarter (October 1 to December 31). It then remained high in following quarters. (The same was true for the number of civil cases reaching a judicial panel for decision.)

cedures, it seems, can alter the outcomes in a wide range of cases concerning other, unrelated areas of law. To set it in a more normative frame: increasing the work of our generalist courts in any one area of law invites spillover effects in other areas of law, if new resources are not also created (and dedicated) to serving those new needs.

To preserve the nature of appellate review, then, new demands put on the courts should be met quickly and flexibly with new judicial resources. If Congress and the President fail to increase judicial resources in keeping with new or growing judicial burdens⁸⁷ — or even if they are just slow to fill judicial vacancies, as is true today⁸⁸ — then the potential costs to the courts may reach beyond those process values commonly said to be at risk (for instance, when a case isn't orally argued, or when a decision doesn't come with reasons in writing).⁸⁹ Rather, the actual outcomes of cases may also be in play.

The story does not end there, however. The prospect of deference drift also raises the possibility of “splits” among the circuits in their intensity of review, due to differences in caseloads. And if the circuits have diverged de facto but not de jure, then oversight through normal means (such as certiorari in the Supreme Court) may be infeasible. This returns the problem again to those who set judicial resources, including the President and Congress, with a further lesson: total judicial resources matter, as already noted — but when uniformity is taken as a further aim, how those total resources are divvied up among jurisdictions also deserves careful attention. As I will explain in Part V, leaving the appeals courts with an uneven distribution of resources (relative to burdens) could generate an “integrity dilemma” — an uneasy bind in which the pursuit of uniformity by conventional means may come at a cost to another judicial priority: the aim of correct results, or what might be called accuracy.

⁸⁷ The judiciary itself also plays a leading role in requesting new judgeships; for a recent appraisal of the practice, see Arthur D. Hellman, *Assessing Judgeship Needs in the Federal Courts of Appeals: Policy Choices and Process Concerns*, 5 J. APP. PRAC. & PROCESS 239 (2003).

⁸⁸ The present delay has been widely noted by both judges and commentators. See, e.g., Letter from Chief Judge Alex Kozinski and Judges of the Ninth Circuit to Senate Leaders (Nov. 15, 2010), available at <http://legaltimes.typepad.com/files/111510-letter-from-9th-circuit.pdf> (noting “our desperate need for judges” and asserting that “we would be greatly assisted if our judicial vacancies — some of which have been open for several years and declared ‘judicial emergencies’ — were to be filled promptly”); see also Dahlia Lithwick & Carl Tobias, *Vacant Stares: Why Don't Americans Worry About How an Understaffed Federal Bench is Hazardous to Their Health?*, SLATE (Sept. 27, 2010), <http://www.slate.com/id/2268466> (analyzing data on the present delay and collecting links to further commentary).

⁸⁹ By no means does this study's focus on case outcomes suggest that process values, including what may be called “procedural justice,” are undeserving of the widespread attention they have received. See Lawrence B. Solum, *Procedural Justice*, 78 S. CAL. L. REV. 181 (2004).

A. “Silent Splits”

Is the intensity of review today, in the two studied circuits, more desirable than what it was before the surge? This is of course a normative call. One might easily argue that it’s just as well to have appellate review with a lighter touch. Or, one might easily argue the opposite. Regardless, there are intriguing problems suggested by these data that do not turn on whether one thinks things better then or now, but depend only on recognizing that the nature of appellate review can vary with the judicial burdens of the moment.

For one, seeing drift across time suggests the possibility of “splits” across circuits. Consider that the Ninth Circuit’s average level of deference to district courts evidently drifted due to the surge. Then, suppose (as is true) that another circuit held steady over the same period. It cannot be the case that these two circuits were matched in their levels of deference, both before and during the surge. Whether they diverged or converged, caseload is demonstrably a key determinant of the gap. Thus, even if the levels of deference in all circuits were to appear steady at a given point in time, one might do well to ask whether there is nonetheless a gap in these steady-state levels, due to preexisting imbalances in judicial burdens.

What is more, such “splits” might emerge (or might already exist) even when the formal doctrines governing review look and sound the same across circuits. This is what I have meant by a “silent split”: the true divergence would not be apparent from reading judicial opinions and their incantations of identical standards of review. Even if two circuits described the abuse-of-discretion standard in the same way, for instance, that doctrine might mean something stricter in practice in a less-burdened circuit, and something more lenient in a circuit that is overworked — just as its meaning might change in any given circuit as judicial burdens fluctuate.

Sometimes articulated doctrinal changes to the formal standards of review do occur. In the Second Circuit, in fact, this may be happening in a prominent category of issues: The standard for reviewing so-called “mixed questions of law and fact” in this circuit had long been described as nondeferential and *de novo*.⁹⁰ As late as 2001, a clearly laid-out taxonomy offered in a Second Circuit opinion unequivocally listed *de novo* as the standard for mixed questions.⁹¹ More recently, however, the standard appears to be giving way, with civil cases start-

⁹⁰ See, e.g., *FDIC v. Providence Coll.*, 115 F.3d 136, 140 (2d Cir. 1997); Evan Tsen Lee, *Principled Decision Making and the Proper Role of Federal Appellate Courts: The Mixed Questions Conflict*, 64 S. CAL. L. REV. 235, 241 (1991) (citing Second Circuit “mixed questions” cases dating from the 1980s).

⁹¹ *Zervos v. Verizon N.Y., Inc.*, 252 F.3d 163, 168 & n.3 (2d Cir. 2001).

ing to invoke an “either/or approach” in which either the (nondeferential) *de novo* or the (deferential) clear error standard may be applied, depending on the question.⁹²

Why does it matter whether the standard formally changes or not? One reason is oversight. When the shift is articulated, as in the example above, a higher authority (such as the circuit court sitting en banc, the Supreme Court, or even Congress) can take notice and ratify, reverse, or modify it. But when caseload-driven drift or “splits” do not show up in articulations of doctrine, and instead occur only in operation, they may escape notice. Moreover, they may resist regulation.

The Supreme Court’s certiorari process, for instance, focuses on seeking out formal “splits” in which the circuits openly adhere to differing doctrines. The search site for disuniformity is the language of judicial opinions. If the Second Circuit says that its standard for reviewing a certain type of ruling is *de novo*, but the Third Circuit says the standard is clear error, then this gap can be closed by certiorari review.⁹³ Such a catch is highly unlikely, by contrast, if these two circuits cite the same standard and yet give it differing force in practice. What it would take to make such a catch is a searching review of decisions in each circuit. This is onerous but not impossible, and maybe it could be encouraged.⁹⁴ (These days, law clerks reading petitions at the Supreme Court are virtually certain to recommend denying the ones that complain of standards of review being misapplied, as matters of mere “splitless fact-bound error correction” and not worth the Court’s time. In theory, the law clerks could be instructed otherwise.)

But such an effort will surely not be enough. There is a reason, after all, for our present scheme of regulating uniformity by standardizing doctrines; the Court simply cannot monitor individual decisions of the circuits, and the threat of (necessarily rare) reprimands can only do

⁹² This “either/or approach,” as applied in civil cases, seems to have come from unlikely origins: a criminal case interpreting a statutory provision for the review of district courts’ applications of sentencing guidelines. *See* *United States v. Vasquez*, 389 F.3d 65, 70 (2d Cir. 2004) (calling it the “either/or approach”). By 2005, however, this option for greater deference had worked its way into the circuit’s civil cases. *See* *Uzdavines v. Weeks Marine, Inc.*, 418 F.3d 138, 143 (2d Cir. 2005). And there it seems to have stayed. *See* *Barszcz v. Dir., Office of Workers’ Comp. Programs & Elec. Boat Corp.*, 486 F.3d 744, 749 (2d Cir. 2007).

⁹³ The task is not as trivial as spotting variations in phrasing, of course, but it is an easy place to start. As Judge Posner recently noted, even though he was “surpris[ed] to find *so* many different standards for awarding attorneys’ fees in Lanham Act cases,” it would not do to assume that they implied more than an illusory split; rather, “[t]o decide whether the standards differ more than semantically would require a close study of the facts of each case.” *Nightingale Home Healthcare, Inc. v. Anodyne Therapy, LLC*, 626 F.3d 958, 962 (7th Cir. 2010).

⁹⁴ In *Gall*, for instance, Justice Stevens concluded that “[a]lthough the Court of Appeals correctly stated that the appropriate standard of review was abuse of discretion, it engaged in an analysis that more closely resembled *de novo* review.” *Gall v. United States*, 128 S. Ct. 586, 600 (2007).

so much. The recourse is to rely on the circuits' self-discipline in staying within the bounds of what a dictated doctrinal formula can sensibly mean. Thus, in the end, it will still fall largely on the overburdened circuit to regulate itself.

What, then, is to be done if "silent splits" are a concern? Consider again the structural remedy of setting resources to meet burdens, and doing so evenly across jurisdictions. Aiming for a resource balance among the circuits may turn out to be the most effective means for enabling (though not ensuring) uniformity. For Congress, this means that in drafting judgeship bills for adding to the federal bench, and in allocating the judiciary's budget, those lawmakers who value uniformity should pay close heed to measurements of per-judge caseloads across the circuits (as adjusted for case difficulty, the ease of recruiting visiting and senior judges to serve on panels, and other such factors).⁹⁵

Attending to balance also means paying mind to the creation of new causes of action or expansions of the federal courts' jurisdiction — and by symmetry, to the elimination of actions or the stripping of jurisdiction. Notably, eliminating diversity jurisdiction has long been proposed by judges and by Congress as a way to lighten caseloads in the federal courts;⁹⁶ but rarely if ever discussed is how doing so may differentially impact the various circuits.

It may be the simplest common sense (though common sense has not always prevailed) that Congress ought to peg the federal courts' resources and personnel to how much work federal law has created for them — and vice versa.⁹⁷ Noticing now the possibility of caseload-

⁹⁵ Concern for balance across circuits has been noted in recent debates over splitting the Ninth Circuit. Senator Dianne Feinstein, for instance, opposed one proposal in 2006 on such grounds (among others):

The split proposal before us would unfairly distribute judicial resources in the West. The Ninth Circuit would keep 71% of the caseload of the current Circuit, but only 58% of its permanent judges. Currently, the Ninth Circuit has a caseload of 570 cases per judge — as opposed to the national average of 381 cases per judge. Under the proposed split, the average caseload in the new Ninth Circuit would actually increase to 600 cases per judge, while the new Twelfth Circuit would have only 326 cases per judge. This inequitable division of resources would leave residents of California and Hawaii facing greater delays, and with court services inferior to their Twelfth Circuit neighbors.

Press Release, Sen. Dianne Feinstein, Senator Dianne Feinstein Cautions Against Splitting Ninth Circuit Court of Appeals for Ideological Motivations (Sept. 20, 2006), *available at* 2006 WLNR 16399308.

⁹⁶ See, e.g., POSNER, *supra* note 7, at 210–21; Friendly, *supra* note 9, at 640–41.

⁹⁷ Just which resources or personnel to add is a harder, and harder-fought, question. What common sense is shared, and what is not, has been laid out by the late Judge Richard Arnold:

Well, if you're not going to have less business, and in fact you're going to have more, and you want to decide the cases promptly, there are only two or three things that can happen to bring that about. The first thing — and this is happening, and this is disturbing — you spend less time on each case. That generally is a bad idea. . . .

The only other solution is more people. . . .

driven “splits,” one might further refine this intuition: if encouraging uniformity of review across jurisdictions is also a priority, then matching resources to burdens is a policy that should be applied not only to the judiciary as a whole, but also circuit by circuit.

B. *Reimagining Boundaries*

Both the urgency and the difficulty of balancing resources (relative to burdens) among the circuits arise from the same source: the very presence of geographic lines. Imagine, then, an approach that seeks to overcome geography by erasing or reshaping boundaries: increasing the flow of judges, court staff, or other resources to where they are most needed at any given time⁹⁸ — or even directing the flow of cases to where the resources are gathered.

As it happens, the surge itself has generated a proposal that would do a bit of both. Judge Newman, formerly chief judge of the Second Circuit and a respected master of judicial management, at the height of the surge reluctantly proposed creating a special court for handling the immigration agency appeals.⁹⁹ Special, but not specialized, this court would be composed of generalist judges on loan from the circuit courts. (This idea is not as radical as it might sound. The present Foreign Intelligence Surveillance Court and the now-defunct Temporary Emergency Court of Appeals are both variants of this model.¹⁰⁰) According to Judge Newman’s proposal, the judges on this special court would be of a “number fluctuating to correspond to the ebb and flow of the [immigration] caseload.”¹⁰¹

This strategy might help relieve differential pressures across jurisdictions by detaching a new judicial burden from its geographic ties.¹⁰²

[But] the proposition that we solve the problem by putting on more staff, it seems to me, ignores the fact that we may be just about at the limit of that solution. . . .
 . . . My feeling is that what we need is more judges. This is very controversial in the federal judiciary and I think I’m in a minority on this subject.

Richard S. Arnold, *The Future of the Federal Courts*, 60 MO. L. REV. 533, 542–43 (1995).

⁹⁸ Today, a version of this is done through an ad hoc and limited visiting judges system. This study’s findings raise the question of whether such a system is enough.

⁹⁹ See *Immigration Litigation Reduction: Hearing Before the S. Comm. on the Judiciary*, 109th Cong. 147–57 (2006) [hereinafter *Immigration Litigation Reduction*] (statement of Jon O. Newman, United States Circuit Judge, United States Court of Appeals for the Second Circuit).

¹⁰⁰ See *Foreign Intelligence Surveillance Court*, FED. JUD. CENTER, http://www.fjc.gov/history/home.nsf/page/courts_special_fisc.html (last visited Jan. 30, 2011); *Temporary Emergency Court of Appeals, 1971–1992*, FED. JUD. CENTER, http://www.fjc.gov/history/home.nsf/page/courts_special_tecoa.html (last visited Jan. 30, 2011).

¹⁰¹ *Immigration Litigation Reduction*, *supra* note 99, at 153 (statement of Jon O. Newman, United States Circuit Judge, United States Court of Appeals for the Second Circuit). Judge Newman made amply clear his general opposition to sending the immigration cases to a centralized court; he proffered this proposal only as a lesser-of-evils option. See *id.*

¹⁰² To imagine the opposite effect, one need look no further than the surge itself. Recall that a jurisdictional rule is what sent such large shares of the immigration surge into the Second and

The spillover effects of the surge would not affect one circuit more than another due to accidents of venue. Yet there is a complication: the scheme could also create differential pressures. The judges staffing this court must come from somewhere; thus, which circuits contributed more judges would become the key factor in determining where the spillover effects would fall. A priority in implementing such a proposal, then, should be to draw help from the lending circuits based on their own per-judge workloads, with the busier circuits contributing fewer judges.

On the spectrum of methods for diverting legal spillovers, creating a new centralized court is at the more controversial extreme. It would also require action by Congress, which may not be forthcoming. (Witness the failure of a Senate proposal in 2006 for redirecting all immigration appeals into the Federal Circuit.¹⁰³) But other methods are possible that could be implemented by the courts alone, and that would shield one aspect of a court's work from overflow in another. Specialized "tracks" within a court's own caseflow are one class of such methods. The Second Circuit's non-argument calendar is a now-familiar example.

One further refinement to such strategies is worth emphasizing: decoupling the source from the solution. Why should it be the immigration agency appeals that are diverted to a special court, or a special track? That they are the cases creating the docket pressure is not a very satisfying answer, as another type of high-volume case might be better suited to specialized treatment than the immigration appeals are. Why not redirect those other cases instead?¹⁰⁴

Or, to see it from another angle, the policy priority might be framed not as redirecting the docket pressure, but as protecting certain cases from its effects. On this view, those other high-priority cases might have the better claim for a separate track. If that sounds sensible, then it should not be surprising that protection is already occurring, with varying degrees of formality. Criminal appeals, for instance, are formally given priority in calendaring in many circuits.¹⁰⁵ Even the daily acts of "triage" in choosing which cases get more (or less) staff or judge time — deciding how much oral argument to allow, if any, or how far

Ninth Circuits — even though all the decisions being reviewed were made by the BIA in a single building in a suburb in northern Virginia.

¹⁰³ See Securing America's Borders Act, S. 2454, 109th Cong. § 501 (2006).

¹⁰⁴ In the case of the Second Circuit, the prevailing wisdom is that the immigration appeals are in fact especially well-suited to the NAC — which is probably a testament to the design of the NAC. Whether other cases are also suitable may soon be seen; the circuit has started to move selected criminal sentencing appeals onto this track as well.

¹⁰⁵ See MCKENNA, HOOPER & CLARK, *supra* note 72, at 51, 104, 117, 143, 156, 171, 185.

a case proceeds along the appellate pipeline — can be said to perform this protective role.

In today's Second Circuit, a leading case on the court's power to dismiss appeals as "frivolous" offers this sharp reminder: "As a matter of policy . . . appellate courts certainly have the inherent authority to allocate scarce judicial resources among the petitions and appeals that press for their attention, and such allocations become especially necessary in this era of burgeoning appellate dockets."¹⁰⁶ Moving certain cases out of bounds, of course, has long been seen as a way of recovering room for the others, and judges have not been short on eloquence in explaining that need.¹⁰⁷

C. *Detecting the Deference*

A sharp reader might observe that even if the circuits were to drift, such changes could nonetheless be detected by measuring reversals — just as this study has done. She might even suggest this practice as a way to improve, or maybe make unnecessary, the pursuit of structural balance. But there is a downside. Consider for a moment the metrics already being widely used as tools for judicial management. Judge Posner, no enemy of empirical measurement, nonetheless warns of the distorting effect entailed by the common circuit practice of using "shame lists" to keep track of which assigned opinions are taking the longest to complete:

The danger of these shaming techniques is that they may give too much salience to what is, after all, only one dimension of judicial performance. If judges are led to think that the world is judging them exclusively on the speed with which they dispatch their business, they will speed up, all right, but the result may be a considerable deterioration in the quality of their decisions.¹⁰⁸

Metrics lend themselves in service of quotas, and spotlighting reversals might well encourage artificial "targeting" of that figure, just as much as for any other metric. A fair response might be that, seeing as the other measures of judicial output are already being used, adding rever-

¹⁰⁶ *Pillay v. INS*, 45 F.3d 14, 17 (2d Cir. 1995).

¹⁰⁷ Most familiar may be Justice Jackson's warning that "[i]t must prejudice the occasional meritorious application to be buried in a flood of worthless ones," for "[h]e who must search a haystack for a needle is likely to end up with the attitude that the needle is not worth the search." *Brown v. Allen*, 344 U.S. 443, 537 (1953) (Jackson, J., concurring in the result). A similar impulse may have fueled Judge Friendly's well-known war against diversity jurisdiction in the federal courts. See Friendly, *supra* note 9, at 641 ("[T]he arguments for retaining [diversity jurisdiction] will not hold water when the federal courts are overburdened with distinctively federal business." (footnote omitted)).

¹⁰⁸ POSNER, *supra* note 7, at 223. In economic terms, this may be seen as a classic "multitask" problem. See Bengt Holmstrom & Paul Milgrom, *Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design*, 7 J.L. ECON. & ORG. 24, 27–28 (1991). I thank Professor Scott Hemphill for calling attention to this parallel.

sal numbers to the mix would only be supplying one further piece of information — maybe balancing out the others. Getting that balance right, though, could hardly be a trivial task.

But there is a deeper problem: this extra piece of information, about reversals, might not be very informative. Recall the reasons deference drift has been difficult to confirm, until now, even by observing reversal rates. The same difficulties apply here: Suppose that one circuit is seen to have higher reversal rates, generally, than another. What of it? Without a deep understanding of the case composition in each circuit, it would be hard to draw credible inferences from that observation. (It could be that the circuit with the higher reversal rate is actually the one reviewing with greater deference — only, it has a higher concentration of reversal-prone cases.) The same goes for a single circuit whose reversal numbers are seen to change over time. This study's empirical contribution is in minimizing such problems, but it has done so only by taking advantage of a fortuitous "natural experiment." This kind of luck is hard to come by.

Making available a richer set of output measures may do some good for guiding a court's priorities, and it may have other benefits such as public notice. But there are risks, and it cannot substitute for the proper distribution of judicial resources. This is especially true for one further reason: neglecting structural balance invites a more profound difficulty — one that can uproot the basic case for consistency and uniformity.¹⁰⁹ This Article concludes with a comment on this destabilizing concern.

V. CONCLUSION: AN INTEGRITY DILEMMA?

Persistent imbalance in judicial burdens across circuits creates a perplexing problem, beyond the inequality of forcing some courts to do more with less, and the consequent risk that litigants in those courts might have a harder time winning an appeal. It also puts the federal courts in a strange bind, in which the forced pursuit of uniformity in appellate review (through such means as oversight or doctrinal control) may come at the expense of another component of a judiciary's integrity: its accuracy.

¹⁰⁹ The standard case for uniformity may be more mantra than gospel; I thank Professor Trevor Morrison for pointing me to very fine new work on reasons we may tolerate or even desire disuniformity in legal interpretation. See Amanda Frost, *Overvaluing Uniformity*, 94 VA. L. REV. 1567, 1571 (2008) (questioning the value of uniformity within the federal judiciary); see also Gil Seinfeld, *The Federal Courts as a Franchise: Rethinking the Justifications for Federal Question Jurisdiction*, 97 CALIF. L. REV. 95, 99–100 (2009) (questioning the mantra of seeking uniformity among the federal judiciary). In thinking about the sort of "splits" I am describing here, though, it helps to distinguish between disuniformity that comes of reasoned disagreement and the kind that occurs by accident or for other unarticulated reasons.

How does such an “integrity dilemma” arise? It comes of mandating a court to do something it lacks the resources to do fully. When an appeals court’s resources are reduced, lightened scrutiny may well be a good thing — a salutary, adaptive form of deference. This may be so, whether the deference is a byproduct of process innovations needed for handling the caseload, or a conscious and sensible reaction to knowing that the court’s relative advantages have declined. Either way, requiring the overburdened appeals court to override this adaptive deference, and instead to exercise more searching review, might in some cases lead to the introduction and not the correction of error.¹¹⁰

Saying whether the data shown here reflect a kind of adaptive deference (and how salutary it might be) will require further, more fine-grained research into the potential mechanisms at work. One aim of this first foray is to motivate follow-up questions such as these: Is the noticed drift due in part to tighter filters on what information gets to the judges or the staff, as a result of more limited briefing, oral argument, or research? Or due to time-shifting of which cases are decided first and which are put off for later?¹¹¹ Or to reassignment in who does what at the court — such as more reliance on staff attorneys, law clerks, or visiting judges? Also generative, it would seem, are deeper inquiries into judicial methods: During the surge, did the affected courts defer more to outside authorities (say, by certifying questions to state courts, or by following the lead of other circuits), just as they did with the district courts?¹¹² Do their opinions show, as Judge Wald might predict, the increased use of limiting devices such as the doctrines of waiver, harmless error, and plain error? If so, were these tools used mainly for making the “easy” cases easier — or for saving the harder questions for a better day?

¹¹⁰ It may seem something of an irony that the dilemma may be lessened if a less munificent theory of deference is credited: that affirming can be tempting under time pressure, as Judge Aldisert cautioned, merely “because a reversal will require a time-consuming, researched opinion.” Aldisert, *supra* note 32, at 43. On such a view, affirmances would be a variant of, rather than a byproduct of, what Judges Newman and Reinhardt have called judicial “shortcuts” and Judge Posner has called “economy measures.” See Posner, *supra* note 40, at 135; Jon O. Newman, *Are 1,000 Federal Judges Enough? Yes. More Would Dilute the Quality*, N.Y. TIMES, May 17, 1993, at A17; Stephen Reinhardt, *Are 1,000 Federal Judges Enough? No. More Cases Should be Heard*, N.Y. TIMES, May 17, 1993, at A17.

¹¹¹ I thank Professor Kate Stith for raising this interesting possibility.

¹¹² Judge Posner has suggested that heavy caseloads contribute to more isolation among the circuits, rather than mutual reference. See *Nightingale Home Healthcare, Inc. v. Anodyne Therapy, LLC*, 626 F.3d 958, 962 (7th Cir. 2010). But might the story differ for issues of first impression?

APPENDIX

I. DATA APPENDIX

The Administrative Office of the U.S. Courts (AO) and the Federal Judicial Center distribute data on all cases filed in the 94 district courts and all appeals filed in the 12 regional appeals courts. These data are used in the federal judiciary's official publications, such as the reports of the Chief Justice and statistics published online. They are made available for researchers through the Interuniversity Consortium for Political and Social Research (ICPSR).

I limit my analysis to variables that are available across all years in the period of study, "statistical years" 1998 to 2005 (fourth quarter of 1997 through third quarter of 2005). As explained in Part III, this window begins where data is first available on whether a pro se litigant was involved in the case. The pro se versus counseled distinction is important for this study, given the special handling of pro se cases by staff attorneys in each appeals court.

In addition, I have limited the sample as follows:

(1) I examine only cases categorized by the AO as "civil" cases. Subject areas within this category include (among others) contracts, torts, property, employment discrimination, other civil rights, intellectual property, regulatory actions, tax, labor laws, social security laws, and commercial law. The asbestos cases are excluded. The "civil" category is distinct from these remaining categories: "criminal," "original writs," and "administrative" (federal agency cases, including the BIA appeals).

(2) I exclude habeas and prisoner suits, even though they are formally categorized as civil cases. As explained in Part III, these cases are especially vulnerable to such shocks as changing case law about the AEDPA or the PLRA.

(3) To avoid double counting, I keep only the record for the "lead" case among consolidated cases. Likewise, I omit reopened appeals and en banc cases; each group is only a miniscule fraction of the caseload.

(4) I omit the D.C. Circuit from all analyses, given its specialized docket. The specialized Federal Circuit is not in the AO appeals data. What remains are all other regional circuits, from First to Eleventh. (Including the D.C. Circuit in the graphs and statistical analyses is immaterial to the results.)

II. ECONOMETRIC APPENDIX

Here, I present measures that quantify the basic outcome changes during the surge period, as seen in the graphs. The table below provides estimates for how much change in the levels of each outcome of

interest is attributable to the surge, as of three years into the surge (second quarter of 2005). This measure is a within-sample prediction. The units are percentage points.

A. Specifications and Robustness

1. *Basic model.* — These measures are based on a regression specification chosen for simplicity and transparency. It is a piecewise linear model, fitting straight lines in each of three periods: before the start of the surge; a “transition” period (to allow for a lag); and a “flooded” period (by which time the surge’s effects, if any, may have started to be felt). In the specification, the dependent variable is the fraction of all terminated appeals that reached a given outcome, within a case category, in a given circuit and in a given quarter¹¹³:

$$\begin{aligned}
 y_{ct} = & \beta_1 \cdot \text{Second} \cdot \text{trend} + \beta_2 \cdot \text{Ninth} \cdot \text{trend} \\
 & + \beta_3 \cdot \text{Second} \cdot \text{transition} + \beta_4 \cdot \text{Second} \cdot \text{change}_t \\
 & + \beta_5 \cdot \text{Second} \cdot \text{flooded} + \beta_6 \cdot \text{Second} \cdot \text{change}_f \\
 & + \beta_7 \cdot \text{Ninth} \cdot \text{transition} + \beta_8 \cdot \text{Ninth} \cdot \text{change}_t \\
 & + \beta_9 \cdot \text{Ninth} \cdot \text{flooded} + \beta_{10} \cdot \text{Ninth} \cdot \text{change}_f \\
 & + \mathbf{X}_{ct} \cdot \mathbf{B}_{11} + \varepsilon_{ct}
 \end{aligned}$$

In this notation, *Second* and *Ninth* are indicators for whether the data point is from the Second or Ninth Circuit; *transition* and *flooded* are indicators for the transition and flooded periods; *c* indexes circuit; *t* indexes quarter of observation; *trend* is time since the start of the data window (capturing a linear trend); and *change_t* and *change_f* measure time since the start of the transition and flooded periods, respectively (capturing the change in trends in those periods). The covariates \mathbf{X}_{ct} include circuit fixed effects, quarter fixed effects, and measures of the “political” composition of the district and appeals judges in a given circuit in a given year.¹¹⁴ The specification includes all components of the interacted terms; components not separately listed are covered by the fixed effects in \mathbf{X}_{ct} .

¹¹³ For instance, a data point for the reversal-rates regression would be the percentage of cases decided, in a given circuit and in a given quarter, that resulted in reversal or remand.

¹¹⁴ The measure is the share of active judges who were appointed by Republican presidents. In the estimations reported here, this measure is included both for the district judges and for the appeals judges, along with a simple interaction term. (Consistent and somewhat sharper results, not reported here, are found using a more compact measure of the chances of having different-party appointees at the district and circuit level.) As might be expected given the difference-in-differences design (which is itself a way of controlling for the changing political composition of the courts during the Clinton and Bush years), and given the fixed effects that are also included, adding (or omitting) these extra control variables does not disturb the finding of falling reversal rates during the surge.

2. *Case-level model.* — The tables show estimates from ordinary least squares (OLS) regressions in which the observations (circuit-by-quarter cell means) are weighted by cell population, and standard errors are clustered by circuit. With this weighting, these circuit-level estimates are essentially equivalent to estimates from a case-level linear probability model (if no case-level covariates are included). The coefficients are identical and standard errors very close between the case-level estimates (not reported here) and the circuit-level estimates (reported here).

3. *Logit model.* — It is also worth noting that a case-level logit specification (not reported here) yields results consistent with the OLS results. The discussion here focuses on the latter.

4. *Treatment of trends.* — In the following table, I show two alternative treatments of trends. From Figure 7, it appears that the reversal rate in the flooded circuits not only declined in absolute terms during the flood, but also turned around what seems to be an upward trend in the years leading up to the surge. Extrapolating that trend, however, requires caution. Because the Second Circuit's vacancy-driven crisis during 1996 through 1998 seems to be the cause of the severe dip in the flooded circuits' combined reversal rate during that period, it may not be ideal to assume that the rising trend in years immediately following would have continued further. I apply two partial solutions to allay this concern. First, in all regressions, I include an indicator for the crisis period in the Second Circuit. Second, as shown in Panel A of the table, I present estimates after removing any contribution of prior trends in the flooded circuits. These estimates are given by $\beta_5 + t \cdot (\beta_6 + \beta_7)$ for the Second Circuit and $\beta_6 + t \cdot (\beta_{10} + \beta_2)$ for the Ninth, where t is the time in quarters between the start of the flooded period and the second quarter of 2005.¹¹⁵ These estimates, in essence, assume a counterfactual in which any such trend would have leveled out at the start of the flooded period. They can thus be thought to represent the simple gap between outcome levels at the start of the surge and at the three-year mark. An alternative treatment is shown in Panel B, in which the contribution of prior trends is included. These estimates are given by $\beta_5 + t \cdot \beta_6$ for the Second Circuit and $\beta_6 + t \cdot \beta_{10}$ for the Ninth. These estimates assume a linear trend in outcomes that would have continued into the surge period. As is apparent from the table, the former approach (shown in the top panel)

¹¹⁵ That is, ten quarters for specifications with the half-year transition period, and eight quarters for those with the full-year transition period. The sum of t and the transition length is always twelve quarters (that is, three years into the surge). The notation above is reduced from $\beta_5 + t \cdot (\beta_6 - (-\beta_7))$ and $\beta_6 + t \cdot (\beta_{10} - (-\beta_2))$, which more clearly shows how the contribution of prior trends is removed.

yields more conservative estimates; I follow that method in presenting further estimates for the subsamples.

5. *Transition period.* — As one further check, I also present all estimates using a separate regression for each of two plausible lengths for the transition period: one half-year and one year.¹¹⁶ These estimates are shown, respectively, as the left and right columns in each set of estimates. As is evident, the estimates are largely insensitive to the chosen length.

B. Basic Estimates

The estimates, as seen in the following table, generally coincide with our visual guesses from “eyeballing” the graphs. As explained above, the numbers represent how much of a change in outcomes (in percentage points) may be attributed to the extra caseload pressure three years into the surge; this is a within-sample prediction. The first row shows estimates based on the full sample of all civil appeals: reversal rates fell approximately four percentage points in each circuit. To put these numbers in context: reversal rates for this full sample averaged approximately ten percent in the Second Circuit and thirteen percent in the Ninth Circuit, from 1999 to 2001.

The table also shows estimates from two subsamples chosen to allay potential concerns. In the second row, labeled “Counseled,” I limit the sample by excluding the pro se cases (those in which at least one of the parties proceeded without an attorney at the time of filing). As previously noted, pro se cases receive specialized treatment from the circuit courts; for instance, they are typically shepherded by dedicated staff attorneys and rarely receive published opinions. Naturally, they also differ in quality from appeals filed by attorneys. Focusing on the “Counseled” subsample avoids the concern that disturbances in the pro se docket may be driving the observed changes in outcomes. In the third row, labeled “Private, counseled,” I further limit the sample to only those counseled cases in which the federal government is not a party. Focusing on these cases avoids the additional concern that many federal government lawyers may have been directly affected by the surge, having been recruited or assigned to work on the immigration case overflow.

Although this analysis has put numbers on the patterns seen in the graphs, a note urging restraint in interpreting these estimates is in or-

¹¹⁶ The tradeoff in defining the transition period is that a longer transition is more likely to cover the true lag (if any), but a shorter one leaves a longer “flooded” period and hence more data for estimating the desired coefficients. It should be emphasized that this transition period need not capture the entire time needed for the surging caseloads (or their effects, if any) to reach their full heights, however; to the contrary, the change-in-trends variables allow for such growth over time.

der: they are specific to the extent and nature of this particular surge, and more needs to be known before these findings can be credibly extrapolated to other events. For easily imagined reasons, it would be premature to conclude that this much increase in a circuit's caseload will always cause that much change in outcomes: cases vary in difficulty, for instance (making it both necessary and challenging to translate docket growth into actual hours of extra work); or there may be tipping points; or differing case compositions might allow differing degrees of "give"; or the circuits might vary in how they absorb fluctuations in caseloads — and this is just for starters.

TABLE. REVERSAL RATES IN CIVIL APPEALS

A. PREDICTED DIVERGENCE AT THREE-YEAR MARK				
<i>Reversal Rates</i>	<i>Second Circuit</i>		<i>Ninth Circuit</i>	
	(1)	(2)	(3)	(4)
<i>All civil appeals</i>	-0.038** [.012]	-0.038** [.014]	-0.039*** [.009]	-0.042*** [.010]
<i>Counseled</i>	-0.048** [.017]	-0.051** [.020]	-0.045*** [.013]	-0.047*** [.014]
<i>Private, counseled</i>	-0.060** [.022]	-0.067** [.026]	-0.036** [.016]	-0.040* [.018]
B. PREDICTED DIVERGENCE FROM TREND AT THREE-YEAR MARK				
<i>All civil appeals</i>	-0.053** [.019]	-0.051** [.020]	-0.054*** [.012]	-0.054*** [.012]

Notes: Estimates represent the predicted divergence of reversal rates, three years into the flood (by second quarter of 2005). They are within-sample predictions. The dependent variables are the percentages of all terminated appeals that are reversed or remanded in a given quarter, in a given circuit (observations are circuit-by-quarter percentages). In columns (1) and (3), the transition period in the model is set at one half-year; in columns (2) and (4), it is set at one year. $N = 32$ quarters \times 11 circuits = 352. All regressions are OLS, weighted by cell population. All regressions include the following controls: quarter fixed effects; circuit fixed effects; measures of the “political” compositions of the district and circuit judges, and their interaction; and an indicator for the vacancy crisis in the Second Circuit. Standard errors, in brackets, are clustered by circuit; *** represents significance at the 1%, ** at the 5%, and * at the 10% levels.