

# Rethinking Judicial Review of High Volume Agency Adjudication

Jonah B. Gelbach & David Marcus\*

*Article III courts annually review thousands of decisions rendered by Social Security Administrative Law Judges, Immigration Judges, and other agency adjudicators who decide large numbers of cases in short periods of time. Federal judges can provide a claim for disability benefits or for immigration relief—the sort of consideration that an agency buckling under the strain of enormous caseloads cannot. Judicial review thus seems to help legitimize systems of high volume agency adjudication. Even so, influential studies rooted in the gritty realities of this decision-making have concluded that the costs of judicial review outweigh whatever benefits the process creates.*

*We argue that the scholarship of high volume agency adjudication has overlooked a critical function that judicial review plays. The large numbers of cases that disability benefits claimants, immigrants, and others file in Article III courts enable federal judges to engage in what we call “problem-oriented oversight.” These judges do not just correct errors made in individual cases or forge legally binding precedent. They also can and do identify entrenched problems of policy administration that afflict agency adjudication. By pressuring agencies to address these problems, Article III courts can help agencies make across-the-board improvements in how they handle their dockets. Problem-oriented oversight significantly strengthens the case for Article III review of high volume agency adjudication.*

*This Article describes and defends problem-oriented oversight through judicial review. We also propose simple approaches to analyzing data from agency appeals that Article III courts can use to improve the oversight they offer. Our argument builds on a several-year study of social security disability benefits adjudication that we conducted on behalf of the Administrative Conference of the United States. The research for this study gave us rare insight into the day-to-day operations of an agency struggling to adjudicate huge numbers of cases quickly and a court system attempting to help this agency improve.*

---

\* Professor of Law, University of Pennsylvania Law School, and Professor of Law, University of Arizona Rogers College of Law. The authors thank BJ Ard, Barbara Atwood, Faisal Chaudry, Andy Coan, Ellie Bublick, Toni Massaro, Nina Rabin, and Christopher Walker for helpful comments on earlier drafts. This Article benefited from a presentation at a September 2017 conference on agency adjudication jointly hosted by the Administrative Conference of the United States and the American Bar Association.

INTRODUCTION .....	1098
I. DISABILITY BENEFITS AND IMMIGRATION ADJUDICATION .....	1103
A. The Exemplar Agencies.....	1103
B. A Brief Primer on the SSA and the EOIR .....	1106
II. THE JUSTIFICATIONS FOR JUDICIAL REVIEW .....	1107
A. The Corrective Function.....	1108
1. <i>The Baseline Problem</i> .....	1108
2. <i>The Costs of Mistakes</i> .....	1112
3. <i>The Opportunity Cost</i> .....	1115
B. The Regulatory and Critical Functions.....	1119
1. <i>The In Terrorem Effect</i> .....	1120
2. <i>The Precedential Effect</i> .....	1121
3. <i>Feedback</i> .....	1125
III. PROBLEM-ORIENTED OVERSIGHT .....	1129
A. Models of Agency Oversight.....	1130
B. The Limits of the Fire Alarm Model .....	1133
C. Problem-Oriented Oversight Through Judicial Review .....	1134
1. <i>Errors</i> .....	1135
2. <i>Problems</i> .....	1136
3. <i>Distinguishing Errors from Problems</i> .....	1140
4. <i>Responding to Problems</i> .....	1145
IV. EVALUATING PROBLEM-ORIENTED OVERSIGHT .....	1148
A. Efficiency.....	1149
1. <i>Resources</i> .....	1149
2. <i>Information Production</i> .....	1151
B. Independence .....	1152
1. <i>Agency Interests</i> .....	1153
2. <i>Political Independence</i> .....	1154
C. Capacity .....	1156
1. <i>Efficacy of Judicial Interventions</i> .....	1156
2. <i>Expertise</i> .....	1158
3. <i>Geographical Dispersion</i> .....	1160
CONCLUSION .....	1161

## Introduction

Federal administrative agencies adjudicate huge numbers of cases. Administrative law judges (ALJs) working for the Social Security Administration (SSA), “probably the largest adjudication agency in the western world,”<sup>1</sup> decided 629,337 claims for disability benefits in 2013.<sup>2</sup> That year, the country’s immigration judges (IJs) completed 253,942 “matters,”<sup>3</sup>

1. *Barnhart v. Thomas*, 540 U.S. 20, 28–29 (2003).

2. SOC. SEC. ADMIN., JUSTIFICATIONS OF ESTIMATES FOR APPROPRIATIONS COMMITTEES FISCAL YEAR 2015, at 144 (2014).

3. U.S. DEP’T OF JUSTICE, EXEC. OFFICE FOR IMMIGRATION REVIEW, FY 2013 STATISTICS YEARBOOK A2 (2014).

and veterans' law judges working for the Board of Veterans Appeals disposed of 41,910 veterans' benefits cases.<sup>4</sup> ALJs at the Office of Medicare Hearings and Appeals issued 79,377 decisions in cases involving Medicare payments and coverage, an effort quickly swamped by the 384,151 new filings the agency received in 2013.<sup>5</sup> Such immense caseloads require agency adjudicators to work with astonishing speed. The average SSA ALJ decided nearly 540 cases in 2013, or more than two per workday,<sup>6</sup> and the average IJ that year resolved matters for more than 1,000 immigrants.<sup>7</sup> The quality of adjudication often buckles under this furious pace, and criticism for slipshod, inconsistent decision-making has long dogged these agencies.<sup>8</sup>

With their power of judicial review, the federal courts sit atop this mountain of adjudication.<sup>9</sup> Time-strapped agency adjudicators have to rule under conditions hardly conducive to thoughtful deliberation. The fact that a federal judge offers a backstop against arbitrary decision-making thus offers something of a psychological salve.<sup>10</sup> Whatever happens within the agency, so the thinking goes, the unfairly denied disability claimant or the immigrant wrongly threatened with deportation can always get justice in an Article III court. For this reason and others, judicial review is thought to "secure an imprimatur of legitimacy for administrative action."<sup>11</sup>

But reality intrudes on this appealing view. The availability of judicial review for what we call "high volume agency adjudication"—adjudication by agencies whose caseloads and available personnel limit adjudicators to no more than a minimal amount of time per case—means that the federal courts feed on a sizable diet of administrative appeals. The 7,225 cases immigrants

4. 2013 BD. OF VETERANS' APPEALS ANN. REP. 24.

5. Statistics are available at Admin. Conference of the U.S. & Stanford Law Sch., *Adjudication Research: Caseload Statistics*, STAN. U., <https://acus.law.stanford.edu/reports/caseload-statistics> [<https://perma.cc/943F-XSRS>].

6. HAROLD J. KRENT & SCOTT MORRIS, STATISTICAL APPENDIX ON ACHIEVING GREATER CONSISTENCY IN SOCIAL SECURITY DISABILITY ADJUDICATION: AN EMPIRICAL STUDY AND SUGGESTED REFORMS 6 (2013); OFFICE OF THE INSPECTOR GEN., SOC. SEC. ADMIN., AUDIT REPORT: THE SOCIAL SECURITY ADMINISTRATION'S EFFORTS TO ELIMINATE THE HEARINGS BACKLOG 4 (2015).

7. FY 2013's 253,942 completed immigration matters were decided by 253 immigration judges on the bench that year. HUMAN RIGHTS FIRST, THE U.S. IMMIGRATION COURT: A BALLOONING BACKLOG THAT REQUIRES ACTION 1 (2016), <https://www.humanrightsfirst.org/resource/us-immigration-court-ballooning-backlog-requires-action> [<https://perma.cc/Z83W-7LC8>].

8. E.g., KRENT & MORRIS, *supra* note 6, at 1–2; Jaya Ramji-Nogales et al., *Refugee Roulette: Disparities in Asylum Adjudication*, 60 STAN. L. REV. 295, 302–03 (2007); James D. Ridgway, *A Benefits System for the Information Age*, 7 VETERANS L. REV. 36, 44 (2015).

9. The federal courts can review agency decisions subject to the limits Congress specifies. *Five Flags Pipe Line Co. v. Dep't of Transp.*, 854 F.2d 1438, 1439 (D.C. Cir. 1988).

10. LOUIS L. JAFFE, JUDICIAL CONTROL OF ADMINISTRATIVE ACTION 320 (1965) (describing judicial review as a "necessary condition, psychologically if not logically, of a system of administrative power which purports to be legitimate").

11. Richard H. Fallon, Jr., *Of Legislative Courts, Administrative Agencies, and Article III*, 101 HARV. L. REV. 915, 942 (1988).

filed in 2013, for instance, accounted for 12.8% of new federal appeals that year.<sup>12</sup> These appeals and others from agencies are indisputably significant to the judicial business of the federal courts.

But is federal court litigation likewise important to harried adjudicators drowning in claims or the agencies that struggle to manage them? The federal courts review only a tiny fraction of the cases agency adjudicators decide—only 3% of SSA ALJ decisions, for example,<sup>13</sup> and only about .03% of decisions by the Office of Medicare Hearings and Appeals.<sup>14</sup> Whatever legitimacy the Article III courts promise must seem like a distant mirage for the vast majority of immigrants, claimants, and others as they litigate in obscure hearing rooms, far away from the grandeur of the federal courts. Doubts that judicial review helps to improve high volume agency adjudication have thus surfaced in administrative law scholarship, perhaps none more importantly than in the seminal studies of social security disability adjudication that Jerry Mashaw wrote in the 1970s and 1980s.<sup>15</sup>

This Article defends the federal courts' involvement in high volume agency adjudication. It has its roots in our sense of what happens day-to-day in hearing offices, immigration courts, and federal judges' chambers around the country. We recently completed a two-year study of social security disability benefits litigation, conducted at the behest of the Administrative Conference of the United States.<sup>16</sup> This study required an extensive quantitative analysis of district court decision-making, as well as scores of interviews with agency officials, ALJs and their support staff, federal judges, and private lawyers. It thus gave us a rich perspective on almost every aspect of federal court involvement with the disability benefits adjudication process. A theoretical companion to the report we produced for the Administrative Conference, this Article uses the trove of information we assembled to inform our understanding of what exactly the federal courts can be—and in some instances are—up to when they review decisions issued by overworked, under-resourced agency adjudicators.

---

12. ADMIN. OFFICE OF THE U.S. COURTS, JUDICIAL BUSINESS 2013, at tbl.B-3 (2013), [http://www.uscourts.gov/sites/default/files/statistics\\_import\\_dir/B03Sep13.pdf](http://www.uscourts.gov/sites/default/files/statistics_import_dir/B03Sep13.pdf) [<https://perma.cc/XT29-DESF>].

13. Nat'l Org. of Soc. Sec. Claimants' Representatives, *Federal Court Filings Increase*, SOC. SEC. F., Aug. 2013, at 14, 14.

14. In FY 2014, claimants filed only twenty-five Medicare appeals in the federal courts. Email from Katherine E. Hosna, Ctrs. for Medicare and Medicaid Servs., to David Marcus, Professor of Law, Univ. of Ariz. Rogers Coll. of Law (May 22, 2017, 12:22 PM) (on file with authors).

15. E.g., JERRY L. MASHAW, BUREAUCRATIC JUSTICE 189–90 (1983); JERRY L. MASHAW ET AL., SOCIAL SECURITY HEARINGS AND APPEALS: A STUDY OF THE SOCIAL SECURITY ADMINISTRATION HEARING SYSTEM 146–47 (1978); Nicholas Bagley, *The Puzzling Presumption of Reviewability*, 127 HARV. L. REV. 1285, 1323 (2014); Paul R. Verkuil & Jeffrey S. Lubbers, *Alternative Approaches to Judicial Review of Social Security Disability Cases*, 55 ADMIN. L. REV. 731, 778, 780 (2003).

16. See JONAH B. GELBACH & DAVID MARCUS, A STUDY OF SOCIAL SECURITY DISABILITY LITIGATION IN THE FEDERAL COURTS (2016).

Our main contribution is to identify a previously unappreciated function that courts perform when they review high volume agency adjudication. Judges correct adjudicators' errors, and they forge precedent to regulate agency decision-making. These jobs are well known, although this Article provides a badly needed reassessment of how well courts tackle them. The function not evident to critics of judicial review is a task we call "problem-oriented oversight." Courts identify and respond to entrenched problems of internal agency administration that can afflict adjudication. When bias discolours an IJ's decision-making and the agency does not respond, for example, courts can do so effectively. When the SSA issues a guidance document that distorts ALJ orders denying disability benefits claims, the federal courts can push the agency to correct course. Problem-oriented oversight involves more than the correction of adjudicator error or the issuance of precedent-setting opinions. The federal courts use various tools at their disposal to hold agencies accountable and insist that they improve. Added to the other functions federal courts discharge, problem-oriented oversight strengthens the case for Article III review of high volume agency adjudication.

Our argument toggles between the descriptive and the normative. Courts presently engage in problem-oriented oversight. We identify the function and describe how federal judges perform it. We also explain how courts can use a straightforward data gathering and analysis method to conduct oversight more rigorously. Finally, we defend the federal courts' oversight capacity. Institutional features of courts and agencies limit how well federal judges can correct adjudicators' errors and regulate agencies through precedent. These impediments pose less of a problem to courts' oversight function. By relying upon a process that requires aggrieved parties to bring problems to their attention, the federal courts can assemble information about poor agency performance efficiently. Their independence from agencies and Congress enables federal judges to address pathologies afflicting agency decision-making without politics or other agency priorities getting in the way. Finally, the federal courts' geographic dispersion and prestige make them effective overseers of a sprawling system of agency adjudication, and the sort of data gathering and analysis problem-oriented oversight requires fit within courts' competencies.

Understanding problem-oriented oversight is important for several reasons. First, appeals from overwhelmed agency adjudicators compose a large chunk of the federal courts' docket. In 2013, for instance, claimants appealed 18,779 SSA ALJ decisions to federal district courts,<sup>17</sup> nearly

---

17. *Appeals to Court as a Percentage of Appealable AC Dispositions*, SOC. SEC. ADMIN., [https://www.ssa.gov/appeals/DataSets/AC04\\_NCC\\_Filed\\_Appealable.html](https://www.ssa.gov/appeals/DataSets/AC04_NCC_Filed_Appealable.html) [https://perma.cc/TS3M-73RE].

equaling federal habeas corpus filings.<sup>18</sup> A fully informed perspective on what Article III judges do on a daily basis requires an appreciation for problem-oriented oversight.

Second, legislators, judges, agency officials, and scholars frequently call for changes to various systems of high volume agency adjudication. Proposals have included the centralization of judicial review in a single Article III court,<sup>19</sup> retrenchment of Article III review,<sup>20</sup> and the end to Article III review altogether.<sup>21</sup> To our minds, problem-oriented oversight, when added to the other functions judges discharge when they oversee high volume agency adjudication, tips an otherwise equivocal normative balance in favor of the current system. But the costs and benefits of judicial review are difficult to measure with precision. Reasonable people may ultimately disagree with our assessment of other functions' efficacy and what problem-oriented oversight adds to the case they present for judicial review. At the least, however, any suggestion to replace Article III review is incomplete unless it grapples with how the change would affect the federal courts' capacity to discharge all of the functions they perform, including problem-oriented oversight.

Third, although courts do engage in problem-oriented oversight, some do so unevenly. In certain instances, federal judges have not yet addressed problems of internal agency administration that need a response. Our description and defense of problem-oriented oversight is an attempt to spur courts to execute this function more evenly and aggressively. Finally, problem-oriented oversight is not something exclusive to high volume agency adjudication. Courts have the capacity to perform this function in any domain where they review large numbers of decisions made by other institutions.<sup>22</sup> An appreciation for problem-oriented oversight and how it works can improve the contributions to good government that generalist judges make in a number of fields.<sup>23</sup>

Part I explains why we use immigration and disability benefits adjudication as the two exemplar systems we draw upon in this Article. It

---

18. JUDICIAL BUSINESS 2013, *supra* note 12, at tbl.C-2A, [http://www.uscourts.gov/sites/default/files/statistics\\_import\\_dir/C02ASep13.pdf](http://www.uscourts.gov/sites/default/files/statistics_import_dir/C02ASep13.pdf) [<https://perma.cc/V4MN-4RRH>].

19. Stephen H. Legomsky, *Restructuring Immigration Adjudication*, 59 DUKE L.J. 1635, 1685–86 (2010).

20. Paul R. Verkuil, *An Outcomes Analysis of Scope of Review Standards*, 44 WM. & MARY L. REV. 679, 725–26 (2002).

21. *Id.* at 728.

22. See, e.g., *Milke v. Ryan*, 711 F.3d 998, 1022–23 (9th Cir. 2013) (Kozinski, C.J., concurring) (blasting the “ridiculous” fact that a police detective with an extensive record of improper conduct was “sent to interrogate a suspect without a tape recorder, a video recorder, a witness or any other objective means of documenting the interrogation”).

23. For an argument that criminal courts should engage in a version of the oversight we describe here, see Andrew Manuel Crespo, *Systemic Facts: Toward Institutional Awareness in Criminal Courts*, 129 HARV. L. REV. 2049, 2052 (2016).

also gives brief introductions to both, to provide basic background for the discussion that follows. Part II includes an extensive assessment of the previously identified functions that the federal courts play when they decide appeals from high volume agency adjudicators. Although our reasons differ, we ultimately agree with Mashaw's influential critique; courts cannot discharge these functions successfully enough to justify the case for Article III involvement in high volume agency adjudication. In Part III, we define problem-oriented oversight and explain how courts engage in it. We also offer a method for data gathering and analysis that courts can use to perform the function more rigorously. Part IV defends problem-oriented oversight through judicial review, stressing the federal courts' institutional advantages as reasons why the task suits them.

## I. Disability Benefits and Immigration Adjudication

### A. *The Exemplar Agencies*

Federal administrative adjudication comes in many varieties. Adjudication by the five ALJs working for the Securities and Exchange Commission represents one variant. They preside over proceedings that often last months and resemble civil litigation in Article III courts.<sup>24</sup> A world apart is a tribunal like the Veterans Administration's Board of Veterans' Appeals. Its sixty-one veterans' law judges decided 41,910 cases in 2013, or 687 per adjudicator.<sup>25</sup> This sort of high volume adjudication poses a distinctive set of challenges. How can large numbers of adjudicators administering the same complex regulatory regime decide cases consistently? How can they render high-quality decisions without allowing a huge backlog of claims to grow? What ensures that adjudicators, worn down by an unending river of cases, do not burn out or become jaded? Finally, can these adjudicators make decisions that will withstand federal judicial scrutiny? Should they be forced to do so?

To assess the contributions federal courts can make to these questions' answers, we draw on the illustrative experiences of the SSA and the Department of Justice's Executive Office for Immigration Review (EOIR). A number of federal agencies engage in high volume adjudication. Table 1 lists those agencies whose hearing-level adjudicators decide more than one case per workday.<sup>26</sup>

---

24. *Bandimere v. SEC*, 844 F.3d 1168, 1170–71 (10th Cir. 2016); Breon S. Peace & Elizabeth Vicens, *Changes and Challenges in the SEC's ALJ Proceedings*, HARV. L. SCH. F. CORP. GOVERNANCE & FIN. REG. (Nov. 12, 2016), <https://corpgov.law.harvard.edu/2016/11/12/changes-and-challenges-in-the-secs-alj-proceedings/> [https://perma.cc/42X9-AML2].

25. All data in this Part on caseloads and numbers of agency adjudicators come from Admin. Conference of the U.S. & Stanford Law Sch., *Adjudication Research: Caseload Statistics*, STAN. U., <https://acus.law.stanford.edu/reports/caseload-statistics> [https://perma.cc/943F-XSRS].

26. By "hearing-level" we mean adjudicators who hold merits hearings to gather evidence, hear from witnesses, and so forth.

Table 1. High Volume Agency Adjudication

Agency Name	Number of Decisions, FY 2013	Number of Agency Adjudicators	Decisions per Adjudicator
Board of Veterans' Appeals	41,910	61	687
Department of Agriculture Administrative Review Branch	1,258	4	314.5
Office of Medicare Hearings and Appeals	79,377	65	1221.2
HHS Provider Reimbursement Board	1,833	5	366.6
EOIR	253,942	248	1024
SSA	793,580	1486	534

We use the EOIR and SSA for several reasons. First, for a long time these agencies have adjudicated more cases than any other.<sup>27</sup> A study of high volume agency adjudication that did not reflect the EOIR's and SSA's experiences with the federal courts would offer narrow instruction. Second, both of these agencies generate significant numbers of federal court appeals. Due to a recent spike, ALJs at the Office of Medicare Hearings and Appeals (OMHA) now decide hundreds of thousands of cases each year. Yet very few of the medical service providers contesting a reimbursement decision ultimately seek judicial review. The federal courts received only twenty-seven appeals from OMHA ALJs in 2016.<sup>28</sup> Likewise, veterans appealed only 109 cases to the Federal Circuit in FY 2015,<sup>29</sup> a year the Board of Veterans'

27. The Office of Medicare Hearings and Appeals now has a caseload roughly equal to EOIR's. This is a recent change, with filings growing 315% between 2010 and 2016. OFFICE OF MEDICARE HEARINGS AND APPEALS, FY 2018 CONGRESSIONAL JUSTIFICATION 7 (2017).

28. Hosna, *supra* note 14.

29. ADMIN. OFFICE OF THE U.S. COURTS, JUDICIAL BUSINESS 2015, at tbl.B-8 (2013), [http://www.uscourts.gov/sites/default/files/data\\_tables/B08Sep15.pdf](http://www.uscourts.gov/sites/default/files/data_tables/B08Sep15.pdf) [https://perma.cc/7HV8-HEZZ].



Appeals received 69,957 cases.<sup>30</sup> In contrast, social security and immigration appeals to the federal courts number in the thousands every year. For an agency like the OMHA, judicial review truly is a mirage. For the SSA and the EOIR, it is a more meaningful component in an overall system of adjudication.

Third, decisions go directly from the SSA and the EOIR to the Article III courts, without some other independent tribunal involved as an intermediary. Before veterans can appeal to the Federal Circuit, they first must litigate before the Court of Appeals for Veterans' Claims (CAVC), an Article I tribunal independent of the Veterans' Administration.<sup>31</sup> Adjudicators at the Internal Revenue Service's Office of Appeals decide more than 40,000 cases each year. Appeals from their orders go almost entirely to the U.S. Tax Court, also an Article I tribunal, before appeals can proceed to a federal appellate court.<sup>32</sup> No such court stands between the EOIR and the courts of appeals, or between the SSA and the district courts, to provide an intermediate level of oversight.

Notwithstanding the agencies' distinctive features, lessons from the EOIR's and SSA's interactions with the federal courts can readily inform critical evaluations of other systems of judicial review. Whether direct oversight by Article III courts succeeds should inform judgments of whether an Article I intermediary works better, for instance. Whether Congress should raise or reduce amount-in-controversy requirements for OMHA appeals, to use another example, should depend at least in part on the desirability of judicial review in Article III courts.<sup>33</sup> Also, much of what can be learned from the interactions between the EOIR and the federal courts, or from those between the SSA and the federal courts, does not depend on the precise configuration of judicial review that these systems' designs involve. The CAVC, for instance, could engage in the sort of data gathering we describe in Part III and use what it assembles to identify and respond to the kind of problems we identify.

---

30. 2015 BD. OF VETERANS' APPEALS ANN. REP. 17; *see also* 2015 U.S. CT. OF APPEALS FOR VETERANS CLAIMS ANN. REP., <https://www.uscourts.cavc.gov/documents/FY2015AnnualReport.pdf> [<https://perma.cc/SDX8-XS2K>].

31. Michael P. Allen, *Due Process and the American Veteran: What the Constitution Can Tell Us About the Veterans' Benefits System*, 80 U. CIN. L. REV. 501, 505 (2011).

32. Leandra Lederman, *Tax Appeal: A Proposal to Make the United States Tax Court More Judicial*, 85 WASH. U. L. REV. 1195, 1196–97 (2008).

33. One of the reasons why so few OMHA decisions get appealed to the federal courts is the amount-in-controversy requirement that federal court jurisdiction over these cases requires. *See* Medicare Appeals Amount in Controversy Threshold Amounts, 81 Fed. Reg. 65,651 (Sept. 23, 2016) (announcing 2017 amount-in-controversy threshold amounts as \$160 for ALJ hearings and \$1,560 for judicial review).

*B. A Brief Primer on the SSA and the EOIR*

The rest of this Article draws upon the EOIR's and the SSA's relationships with the federal courts to inform our claims about judicial review and the functions it plays in the context of high volume agency adjudication. Both systems have endless complexities, but a basic orientation to each should suffice for what follows.

As of June 2017, the EOIR, part of the U.S. Department of Justice (DOJ), employed about 325 IJs who work in dozens of immigration courts scattered around the country.<sup>34</sup> Cases can get before IJs in several ways. An immigrant who claims to be fleeing persecution can apply for asylum with the U.S. Citizenship and Immigration Services.<sup>35</sup> If USCIS rejects her application, it will forward her case to an IJ for an asylum hearing.<sup>36</sup> Alternatively, the government might initiate removal proceedings against an undocumented immigrant picked up at a work site, or against a noncitizen arrested for a crime. These cases go directly to IJs for adjudication. The IJ holds a hearing and issues a decision on the immigrant's asylum petition or request for cancellation of removal.<sup>37</sup> If the immigrant loses, she can ask the Board of Immigration Appeals (BIA), a sixteen-member appellate tribunal located at EOIR's headquarters in Falls Church, Virginia, to review the IJ's decision.<sup>38</sup> The immigrant can appeal from an adverse BIA decision to "the court of appeals for the judicial circuit in which the immigration judge completed the proceedings."<sup>39</sup>

The SSA's Offices of Hearings Operations and of Analytics, Review, and Oversight encompass an enormous system of disability benefits adjudication. A person who believes that his impairments prevent him from working applies for disability benefits at one of the SSA's 1,300 field offices.<sup>40</sup> If initially denied, and if denied again upon reconsideration, the claimant can request a hearing before an ALJ.<sup>41</sup> (From this point on, "ALJ"

---

34. Press Release, Exec. Office for Immigration Review, Executive Office for Immigration Review Swears in 11 Immigration Judges (June 16, 2017), <https://www.justice.gov/eoir/pr/executive-office-immigration-review-swears-11-immigration-judges> [<https://perma.cc/Q4EP-JHB3>]. As of the time of writing, the EOIR is expanding the IJ corps considerably. Statement of James McHenry, Acting Director, Exec. Office for Immigration Review, Before the Subcommittee on Immigration and Border Security, Committee of the Judiciary, U.S. House of Rep., Nov. 1, 2017, at 3.

35. *Obtaining Asylum in the United States*, U.S. CITIZENSHIP & IMMIGRATION SERVS., <https://www.uscis.gov/humanitarian/refugees-asylum/asylum/obtaining-asylum-united-states> [<https://perma.cc/G544-VUJF>].

36. *Id.*

37. 8 U.S.C. § 1229(a) (2012).

38. 8 C.F.R. § 1003.1 (2017); *Board of Immigration Appeals*, U.S. DEP'T JUSTICE, <https://www.justice.gov/eoir/board-of-immigration-appeals> [<https://perma.cc/8ZQQ-VG4V>].

39. 8 U.S.C. § 1252(b)(2).

40. GELBACH & MARCUS, *supra* note 16, at 16–17.

41. *See id.* at 17–18.

refers to an SSA ALJ.) The ALJ works with about 1,400 judicial colleagues in one of 160 hearing offices around the country.<sup>42</sup> Aided by a “decision writer,” the SSA’s version of a law clerk, the ALJ issues a written decision after considering the claimant’s medical records, his hearing testimony, and other evidence.<sup>43</sup> If the decision goes against the claimant, he can appeal to the SSA’s Appeals Council, located in the same nondescript Falls Church office building. After a workup by an “analyst,” who also functions as a law clerk, the case goes to one of dozens of appellate adjudicators for a decision.<sup>44</sup> If the claimant loses again, he can appeal to a federal district court, typically the one in the district where he resides.<sup>45</sup>

## II. The Justifications for Judicial Review

Disability benefits adjudication belongs as an exemplar in a study of judicial review in part because it has attracted the most exhaustive attention. No treatment of SSA decision-making is more important than the landmark report Mashaw and his colleagues compiled in 1978. They identified several possible functions that judicial review performs, including the following:

- A “corrective function”: courts can correct erroneous agency decisions.
- A “regulative function”: courts can induce agency adjudicators to decide cases more accurately, either through fear of judicial reversal (“the in terrorem effect”) or by forcing them to abide by court-fashioned rules (“the precedential effect”).
- A “legitimizing function”: review of an agency’s decision by an independent judiciary can increase public confidence in the legitimacy of outcomes.
- A “critical function”: courts offer agencies a “steady stream” of feedback that they can use to improve, and that is valuable for its own sake.
- A “public information function”: court decisions “serve as a window on an agency whose operations would otherwise be largely invisible.”<sup>46</sup>

Primarily assessing the corrective and regulative functions, the Mashaw group concluded that judicial review’s benefits for the adjudication of social security disability claims did not justify its costs.<sup>47</sup> Decades later, this claim

---

42. *Id.*

43. *Id.* at 20–23.

44. *Id.* at 27–28.

45. 42 U.S.C. § 405(g) (2016); *see also* GELBACH & MARCUS, *supra* note 16, at 30–35.

46. MASHAW ET AL., *supra* note 15, at 136–37.

47. *Id.* at 146–47.

continues to reverberate in discussions of whether the federal courts should review agency adjudication.<sup>48</sup>

The Mashaw group's discussion remains the most comprehensive and trenchant analysis of judicial review of high volume agency adjudication. It thus offers a good template for an inquiry into what functions judicial review can serve and how well it can perform them. Revisited four decades later, much of the Mashaw group's skepticism remains warranted, and not just for disability benefits adjudication. What follows updates and elaborates on the Mashaw group's analysis, with a focus on judicial review's error correction, regulative, and critical functions.<sup>49</sup> In any odd instance, the federal courts can discharge one or more of these functions well. But institutional features of courts and agencies prompt doubts that the former can do so reliably enough to place judicial review of high volume agency adjudication on stable normative footing.

#### A. *The Corrective Function*

Plenty of appeals filed in the federal courts involve mistakes made by agency adjudicators. To think otherwise requires unwarranted confidence in the internal agency appellate tribunals that stand between first-line adjudicators and the federal courts. Year after year, the SSA requests a voluntary remand in about 15% of cases appealed to the federal courts.<sup>50</sup> These "RVRs" happen only when an SSA lawyer and the Appeals Council conclude that the lawyer cannot defend the ALJ's decision as compliant with the agency's own view of social security law and policy.<sup>51</sup> Disability appeals go to the federal courts only after Appeals Council review, so RVRs amount to a concession that internal appellate review sometimes fails.

Errors surely remain for the federal courts to correct, and federal courts surely correct errors. But the Mashaw group doubted that courts can do so reliably. We disagree. Nonetheless, the opportunity cost of court-based error correction unsettles its contribution to the case for judicial review.

*1. The Baseline Problem.*—The Mashaw group questioned the capacity of courts to correct errors because of doubts that judges could evaluate

---

48. *E.g.*, Bagley, *supra* note 15, at 1330.

49. Following the Mashaw group's lead, we do not assess the legitimizing and public information functions at any length. Given the small numbers of claimants and immigrants who avail themselves of federal court review, we doubt that, for the average person caught up in high volume agency adjudication, the distant prospect of judicial review meaningfully legitimizes the exercise of agency adjudicator power. *See* MASHAW ET AL., *supra* note 15, at 147 (discounting the legitimizing function for a similar reason). Courts do broadcast information about agency adjudication that might not otherwise surface, but a judgment about the value of judicial review should account for the type of information publicized. When courts engage in problem-oriented oversight, they bring to light information germane to a critical evaluation of agency adjudication.

50. GELBACH & MARCUS, *supra* note 16, at 31.

51. *Id.* at 32.

disability claims as accurately as ALJs.<sup>52</sup> The problem involves a contrast between courts' and ALJs' baselines. ALJs handle a much larger caseload than federal judges, and ALJs get their cases earlier in the adjudication process. ALJs thus see a wider array of types of claims than federal judges do. Moreover, the government cannot appeal, so claimants pick all of the cases that go to federal court.<sup>53</sup> An ALJ may therefore have a different "cutpoint"<sup>54</sup>—roughly, the line the ALJ would draw along a given dimension between disability and no disability—than a federal court for a decision in favor of the claimant. Figures 1 and 2 illustrate.

Figure 1. ALJ Baseline

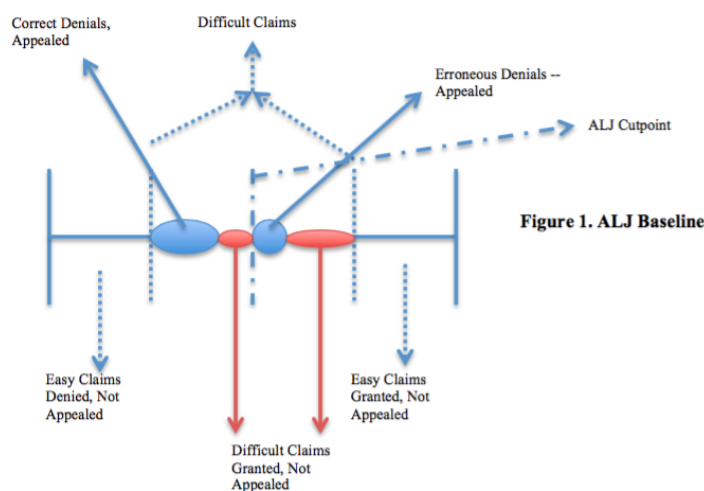
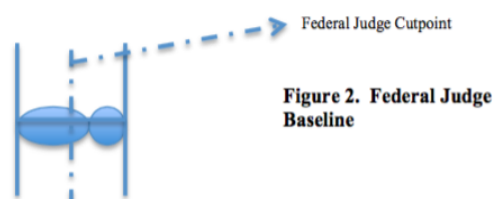


Figure 2. Federal Judge Baseline



Most appeals presumably come from the groups of correct and erroneous denials of what we call "difficult claims." Bereft of a more diverse baseline, a federal judge might view what to the ALJ was a relatively weak

52. MASHAW ET AL., *supra* note 15, at 138–39.

53. For analogous information about immigration appeals, see Stephen H. Legomsky, *Restructuring Immigration Adjudication*, 59 DUKE L.J. 1635, 1649 n.64 (2010).

54. Jonathan Masur, *Patent Inflation*, 121 YALE L.J. 470, 483 (2011).

claim for benefits as an above-average one.<sup>55</sup> “If federal judges saw more of what ALJs grant,” this ALJ told us, “they would appreciate why a case seems more borderline to an ALJ.”<sup>56</sup>

The baseline problem can manifest itself in more granular ways. A federal judge might react differently than an agency adjudicator to particular evidence, for instance. With their immense caseloads, ALJs and decision writers can see letters from the same physicians that use the same phrases to describe patients with strikingly similar problems.<sup>57</sup> “We know which doctors are trustworthy and which ones aren’t,” one ALJ told us, “but we can’t put this in a decision.”<sup>58</sup> Likewise, another ALJ said, “claimants can testify in an obviously coached manner, taught to say just the right thing to buttress a claim for benefits.”<sup>59</sup> IJs may experience the same phenomenon.<sup>60</sup> An ALJ or IJ might correctly discount such evidence, but a federal judge with a narrower evidentiary baseline might fault the ALJ for doing so.

Federal judges have countervailing institutional advantages, however, that may exceed whatever edge a richer baseline gives ALJs. Perhaps most importantly, courts can invest more time and resources in decision-making than agency adjudicators can. To keep backlogs at bay, the SSA asks its ALJs to decide between 500–700 cases per year,<sup>61</sup> with each involving hundreds of pages of medical records and a complex regulatory regime. This caseload is “preposterous,” as one district judge described it.<sup>62</sup> ALJs spend about two-and-a-half hours total on all aspects of a case, and decision writers an additional eight hours when drafting a decision denying a claim.<sup>63</sup> A case gets about four hours of analyst time at the SSA’s Appeals Council, and appellate adjudicators decide five to twelve cases per day.<sup>64</sup>

---

55. MASHAW ET AL., *supra* note 15, at 139.

56. GELBACH & MARCUS, *supra* note 16, at 77.

57. *Id.* at 77–78.

58. *Id.* at 78; *see also* Lester v. Chater, 81 F.3d 821, 832 (9th Cir. 1995) (disapproving of the ALJ’s “unsupported and unwarranted speculation that the . . . doctors were misrepresenting the claimant’s condition or were not qualified to evaluate it”).

59. GELBACH & MARCUS, *supra* note 16, at 78.

60. *See* Jeff Chorney, 9th Cir. Slaps “Incomprehensible Ruling”, NAT’L L.J. (Mar. 21, 2005) (quoting an immigration judge as insisting that arguments from asylum applicants “were all the same”).

61. GELBACH & MARCUS, *supra* note 16, at 36.

62. *Id.* at 73 n.404; *see also* Alex M. Parker, *Recession Is Exacerbating Social Security Claims Backlog, Panelists Say*, GOV’T EXEC. (May 28, 2009), [www.govexec.com/oversight/2009/05/recession-is-exacerbating-social-security-claims-backlog-panelists-say/29262/](http://www.govexec.com/oversight/2009/05/recession-is-exacerbating-social-security-claims-backlog-panelists-say/29262/) [<https://perma.cc/3493-TRB7>] (quoting a federal magistrate judge describing ALJ workloads as “unconscionable”).

63. GELBACH & MARCUS, *supra* note 16, at 14, 24.

64. *Id.* at 29.

With 1,000 cases to decide each year, IJs face an even more herculean task.<sup>65</sup> BIA review practices have changed considerably over the last fifteen years, but at their nadir, caseloads gave board members only 7–10 minutes for the average case.<sup>66</sup> Federal judges have more time to deliberate.<sup>67</sup> In FY 2014, when on average a single IJ had more than 1,400 matters on his docket,<sup>68</sup> the entire federal appellate bench received 54,988 filings.<sup>69</sup> Given the governing law's endless details and the often sizable case files assembled before agency adjudicators, the sheer amount of time a federal judge might spend compared to an ALJ or IJ can compensate for the narrower baseline.

Another institutional advantage adds to the courts' side of the ledger. The decision-writer-to-ALJ ratio is 1:1,<sup>70</sup> for instance, and the law-clerk-to-IJ ratio is 1:4.<sup>71</sup> District judges have at least two clerks, and court of appeals judges typically have four.

Agency adjudicators' baselines may give them a better sense of the overall disability landscape than what federal judges enjoy. But the time and resource shortfalls that afflict agency decision-making may make its adjudicators more error-prone, while federal judges' comparative surfeit of both improves their relative capacity to decide cases accurately. How these advantages and disadvantages balance out is not obvious in the abstract. Not long ago, however, the SSA's Chief ALJ conceded that it favors the federal courts, observing that "most of our decisions that are remanded or reversed by the federal judges are remanded or reversed simply because our decision did not comply with our own policy."<sup>72</sup> Although the SSA has embarked upon an extensive program of quality improvement since these comments, the composition of the pool of federal court appeals probably has not changed all that much since that time, as we argue at length in our report.<sup>73</sup> Federal judges can probably identify flawed decisions fairly accurately. The same is

65. HUMAN RIGHTS FIRST, REDUCING THE IMMIGRATION COURT BACKLOG AND DELAYS 5 (2016). See generally U.S. DEP'T OF JUSTICE, OFFICE OF THE INSPECTOR GEN., I-2013-001, MANAGEMENT OF IMMIGRATION CASES AND APPEALS BY THE EXECUTIVE OFFICE FOR IMMIGRATION REVIEW (2012) (documenting flaws in EOIR processes and recommending reforms).

66. *Kadia v. Gonzales*, 501 F.3d 817, 820 (7th Cir. 2007).

67. GELBACH & MARCUS, *supra* note 16, at 73.

68. *Empty Benches: Underfunding of Immigration Courts Undermines Justice*, AM. IMMIGR. COUNCIL (June 17, 2016), <https://www.americanimmigrationcouncil.org/research/empty-benches-underfunding-immigration-courts-undermines-justice> [<https://perma.cc/Q2X6-AXVF>].

69. ADMIN. OFFICE OF THE U.S. COURTS, JUDICIAL BUSINESS 2014, at tbl.1, [www.uscourts.gov/statistics-reports/us-courts-appeals-judicial-business-2014](http://www.uscourts.gov/statistics-reports/us-courts-appeals-judicial-business-2014) [<https://perma.cc/4D8N-T2ES>]. During a several-year period in the mid-2000s, the Second and Ninth Circuits lost their time advantage over IJs and the BIA for immigration cases. But a return of immigration appeals to lower levels has restored it.

70. GELBACH & MARCUS, *supra* note 16, at 74.

71. HUMAN RIGHTS FIRST, *supra* note 65, at 5 n.41.

72. Letter from Frank A. Cristaudo, Chief Admin. Law Judge, Office of Disability Adjudication and Review, to Colleagues 3 (Dec. 19, 2007) (on file with authors).

73. GELBACH & MARCUS, *supra* note 16, at 54.

likely true of immigration appeals, at least for the cases that the federal courts remand to the agency.<sup>74</sup>

2. *The Costs of Mistakes.*—Whatever the frequency, surely federal judges err and incorrectly remand cases from time to time. The error-correction function cannot justify judicial review if judges make costly mistakes. Suppose a judge is right eight times out of ten when she remands a case to the agency. Judicial review would prove harmful on balance if the costs of the false positives (the two erroneous remands) exceed the benefits of the true positives (the eight correct ones).

The cost–benefit balance resists an easy assessment in part because the social value and harms of wrongfully made disability payments and of payments wrongfully withheld cannot really be measured.<sup>75</sup> One estimate holds that the wrongful allowance of benefits from 2005–2014 will ultimately cost the federal treasury \$72 billion.<sup>76</sup> On the other side of the ledger is an actually disabled claimant whose impairments make a correct decision on her claim “a matter of life and death.”<sup>77</sup> How does the social value of a true positive compare to the costs of false positives?

Any estimate of this balance must necessarily be crude. But one guess suggests that the benefits of true positives basically equal the costs of false positives in the aggregate, at least for social security adjudication, where the likelihood and costs of false positives relative to other categories of high volume agency adjudication are highest.<sup>78</sup> A claimant who successfully

---

74. The federal courts remand many fewer immigration cases, percentage-wise, than social security appeals. *See infra* notes 102–03 and accompanying text. Their prerogative to review IJ decisions is very narrow. Jill E. Family, *Threats to the Future of the Immigration Class Action*, 27 WASH. U. J.L. & POL’Y 71, 82–83 (2008). We thus presume that, when a federal court remands an immigration court decision, the likelihood that it is indeed flawed is very high.

75. Jerry L. Mashaw, *The Supreme Court’s Due Process Calculus for Administrative Adjudication in Mathews v. Eldridge: Three Factors in Search of a Theory of Value*, 44 U. CHI. L. REV. 28, 48 (1976).

76. Mark J. Warshawsky & Ross A. Marchand, *Reforming the System of Review by Administrative Law Judges in Disability Insurance* 15–16 (Mercatus Ctr., George Mason Univ., Working Paper, 2015), <https://www.mercatus.org/system/files/Warshawsky-Reforming-DI-Review.pdf> [<https://perma.cc/M9J2-UXEC>].

77. GELBACH & MARCUS, *supra* note 16, at 15 (quoting a claimant representative).

78. While we lack equivalent numbers, we are confident that this balance comes out in favor of judicial review for immigration adjudication as well. For one thing, the federal courts rule very infrequently in favor of immigrants. The circuits upheld agency decisions in 88.7% of cases in 2016. John Guendelsberger, *Circuit Court Decisions for December 2015 and Calendar Year 2015 Totals*, IMMIGR. L. ADVISOR, Jan. 2016, at 5, 5. Given these numbers, the 11.3% of cases immigrants win should involve fairly egregious agency errors. Moreover, the harms that result from a false positive—a decision reversing the BIA when the immigrant should be removed or denied asylum—should be fairly low. Immigrants who are ordered deported based on their criminal activity almost never prevail on appeal. Only an immigrant who has no criminal record, and thus presents no indication of a threat to public safety, is likely to prevail erroneously.



obtains benefits can expect to receive about \$1,500 in cash per month.<sup>79</sup> In 2007, the Government Accountability Office determined that SSA ALJs eventually grant benefits to 66% of claimants who secure a court remand.<sup>80</sup> We used these numbers together with a range of assumptions about benefits wrongly provided, the costs associated with ALJ time spent on court remands,<sup>81</sup> the social value of dollars received by disability beneficiaries, and the social costs of raising the tax revenue needed to pay for benefits and the operation of the judicial review system, to conduct a back-of-the-envelope cost–benefit analysis. Our calculations yield two key conclusions. First, using what we regard as reasonable values of the key normative and positive parameters, we find that the net social value of judicial review of disability appeals is likely within \$10–15 million of zero. Second, even with extreme assumptions in either direction, the net social value or cost of judicial review seems very unlikely to be more than a drop in the bucket when measured relative to the overall magnitude of disability (and federal court) expenditures—almost surely less than roughly a tenth of a penny for every dollar spent on these programs.<sup>82</sup>

---

79. This figure refers to a claimant seeking SSDI benefits, not SSI benefits. *Social Security Administration Oversight: Examining the Integrity of the Disability Determination Appeals Process, Part II: Hearing Before the H. Comm. on Oversight and Gov't Reform*, 113th Cong. 50 (2014) (statement of Carolyn Colvin, Acting Comm'r, Social Security Administration). This figure does not include the value of Medicare coverage that a beneficiary would also receive.

80. U.S. GOV'T ACCOUNTABILITY OFF., GAO-07-331, REPORT TO CONGRESSIONAL REQUESTERS, DISABILITY PROGRAMS: SSA HAS TAKEN STEPS TO ADDRESS CONFLICTING COURT DECISIONS, BUT NEEDS TO MANAGE DATA BETTER ON THE INCREASING NUMBER OF COURT REMANDS 16 (2007).

81. In FY 2015, federal courts remanded 8,646 cases to the SSA. *Court Remands as a Percentage of New Court Cases Filed*, SOC. SEC. ADMIN., [https://www.ssa.gov/appeals/DataSets/AC05\\_Court\\_Remands\\_NCC\\_Filed.html](https://www.ssa.gov/appeals/DataSets/AC05_Court_Remands_NCC_Filed.html) [<https://perma.cc/H876-F8BG>]. We use this data and the following assumptions: (1) One-fourth of the 66% of court remands that result in the payment of benefits do so because ALJs want to get rid of troublesome cases, not because the claimant is actually disabled; (2) one-half of the 34% of court remands that do not result in the payment of benefits fail because the federal judge erred, with the other half of remands that do not result in benefit payment being true negatives, i.e., correct denials of benefits; and (3) court remands are more difficult than cases heard in the first instance, such that an average ALJ could decide a dozen new cases during the time required to decide court remands. See GELBACH & MARCUS, *supra* note 16, at 48 & n.291.

82. Here we define the benefits as benefits paid to claimants who should receive them. In FY 2015, courts remanded 8,646 cases to the agency. Assuming that ALJs paid benefits to 49.5% of these claimants correctly (three-fourths of the 66% of claimants who won benefits), judicial review creates an annual benefit of \$77,035,860 in benefits rightly paid to people with disabilities. We let the social value of paying a dollar in benefits to an eligible claimant be  $\alpha$  dollars. For example, if  $\alpha = 2$ , then the social value of providing a dollar to an eligible beneficiary is as good as providing two dollars to a randomly drawn member of the remainder of the population. Thus, the benefit side of having judicial review is \$77,035,860 times  $\alpha$ .

The costs of judicial review include ALJ resources that have to be spent on court remands as well as those federal judicial resources spent handling disability appeals. As far as ALJ resources go, each court remand displaces two cases an ALJ could decide in the first instance. Thus, the 8,646 remands from federal court in FY 2015 displaced 17,292 first-instance remands. In FY 2015, 1,519 ALJs, *Administrative Law Judge—Federal Salaries of 2015*, FEDERALPAY.ORG, <https://www>

.federalpay.org/employees/occupations/administrative-law-judge/2015 [https://perma.cc/B74T-GJL8], decided 507,883 cases, for an average of roughly 334 cases decided per ALJ. 2017 SOC. SEC. ADMIN. BUDGET JUSTIFICATION, LIMITATION ON ADMIN. EXPENSES 75 tbl.3.34 (2016), <https://www.ssa.gov/budget/FY17Files/2017LAE.pdf> [https://perma.cc/3YKV-YAQ7]. At that rate, it would take roughly 52 ALJs to decide 17,292 first-instance cases. In 2015, the average ALJ's salary was \$159,196.65. *Administrative Law Judge—Federal Salaries of 2015*, FEDERALPAY.ORG, <https://www.federalpay.org/employees/occupations/administrative-law-judge/2015> [https://perma.cc/B74T-GJL8]. That year, the SSA spent about 27% of ALJ salaries on fringe benefits. 2017 SOC. SEC. ADMIN. BUDGET JUSTIFICATION, *supra*, at 75 tbl.3.28, <https://www.ssa.gov/budget/FY17Files/2017LAE.pdf> [https://perma.cc/3YKV-YAQ7] (showing that ALJ benefit and salary expenses totaled \$63,610,135 and \$232,875,700, for a ratio of approximately 0.27). Thus the total cost to the SSA in 2015 of court remands, measured in terms of ALJ productivity, is  $52 \times 1.27 \times \$159,196.65$ , which amounts to \$10,510,705.17. Assuming that the cost of decision writers and other support staff for the 52 ALJ-equivalents would amount to another 50% of this figure yields a total SSA staff cost of  $1.5 \times \$10,510,705.17$ , or \$15,766,058.

With respect to judicial resources, the 19,222 disability cases terminated in the twelve months ending June 30, 2015, amounted to 7% of civil terminations. *See* JUDICIAL BUSINESS 2015, *supra* note 29, at tbl.C-4, [http://www.uscourts.gov/sites/default/files/c04jun15\\_0.pdf](http://www.uscourts.gov/sites/default/files/c04jun15_0.pdf) [https://perma.cc/3LXS-S2J3] (counting cases in the Disability Insurance and Supplemental Security Income rows). Assuming these cases would require 7% of the work time of 630 district court judges (663 permanent authorized and 10 temporary authorized, ADMIN. OFFICE OF THE U.S. COURTS, AUTHORIZED JUDGESHIPS, <http://www.uscourts.gov/sites/default/files/allauth.pdf> [https://perma.cc/HGL3-P8YQ]), less 43 vacancies, *Vacancy Summary for June 2015*, ADMIN. OFFICE OF THE U.S. COURTS, <http://www.uscourts.gov/judges-judgeships/judicial-vacancies/archive-judicial-vacancies/2015/06/summary> [https://perma.cc/EW3W-5KRK]), which is high since these cases don't go to trial or involve intensive pretrial wrangling, these cases account for the work time of roughly 44 federal judges.

According to the Administrative Office of the U.S. Courts' *FY 2017 Congressional Budget Summary*, filling an Article III judgeship costs \$233,333.33, *see* ADMIN. OFFICE OF THE U.S. COURTS, FY 2017 CONGRESSIONAL BUDGET SUMMARY 24 (2016), [http://www.uscourts.gov/sites/default/files/fy\\_2017\\_federal\\_judiciary\\_congressional\\_budget\\_summary\\_0.pdf](http://www.uscourts.gov/sites/default/files/fy_2017_federal_judiciary_congressional_budget_summary_0.pdf) [https://perma.cc/9GCR-GWYU] (requesting \$1.4 million to fill six judgeships), plus an additional \$140,000 for each of five staff members, *see id.* at 25 (requesting \$4.2 million for thirty associated staffers). The total including support staff is thus \$933,333.33 per judge or \$41,066,667 for the 44 additional federal judges. Adding that figure to the SSA staff cost of \$15,766,058 calculated above yields a total government staff cost of \$56,832,725. In addition, the SSA must pay some of claimants' litigation costs under the EAJA; in 2015 these costs amounted to \$38,132,381, *Social Security Administration Data for Equal Access to Justice Act Payments*, SOC. SEC. ADMIN., <https://www.ssa.gov/open/data/EAJA.xlsx> [https://perma.cc/R4UM-GMV6], so the total government staff cost and fee-shifting expenses come to \$94,965,106. Government staff must be paid out of tax revenues. Because taxes affect behavior, there are social costs of raising a dollar of tax revenue. *See* JONATHAN GRUBER, PUBLIC FINANCE AND PUBLIC POLICY 566–75 (5th ed. 2016) (cataloguing a variety of tax policies and their varied effects on behavior). An implication is that in general, the overall social cost of having the government spend an additional dollar exceeds one dollar. (This implication might not hold true during severe recessions, a special case.) To account for this issue, we let  $\beta$  be the social cost of raising a dollar of tax revenue, so that under our assumptions the total social cost of government staff work related to judicial review of disability appeals is  $\$94,965,106 \times \beta$ . So far we have a total social value of  $\$77,035,860 \times \alpha$  and a total social cost associated with government staffing equal to  $\$94,965,106 \times \beta$ . The difference will be marginally positive if  $\alpha \geq 1.24 \times \beta$ , i.e., if the social value of transferring a dollar to persons truly entitled to receive disability benefits is roughly 1.24 times the marginal cost of raising a dollar in taxes. We think this assumption is reasonable, though of course the value of  $\alpha$  is fundamentally a normative question.

There is also the question of how to account for benefits erroneously paid to those not actually entitled to them under the law. Under our assumption above, benefits would have been wrongly

3. *The Opportunity Cost.*—Another way to look at error correction is to consider whether the resources it consumes could be spent in alternative ways. On this view, the limitations of the error correction function lie not only with the difficulties judges have identifying errors, nor only with the harms that false positives cause, but rather with judicial review's opportunity cost. If invested in agency adjudication, the resources that judicial review requires might lead to fewer errors made in the first instance.

Any adjudication system should prefer error avoidance to error correction, all else equal. An acquittal or dismissal obviously compares favorably to a conviction that later gets vacated on appeal. If the system's designer has \$100 to spend, and if that sum can either avoid one error or correct one error, the designer should invest in error avoidance rather than error correction. Judicial review makes sense from this perspective only if the \$100 can buy more error correction than error avoidance.

For social security claims, the return on investment probably comes out in favor of error avoidance rather than error correction. At a minimum, resources expended on judicial review include salaries for the SSA litigators who brief and argue cases, Equal Access to Justice Act (EAJA) fees paid to

---

paid to 16.5% (one-quarter of 66%) of the 8,646 claimants who won remands in FY 2015, which amounts to \$25,678,620 in benefits wrongly paid. One approach would be to regard these paid-out benefits as having a net social cost of  $\$25,678,620 \times \beta$ , since taxes must be raised to fund these benefits. But that approach fails to recognize that (i) these benefits have some value to those who receive them, and (ii) the well-being of such recipients has some social value. Presumably the social value of transferring a dollar in disability benefits to those not actually entitled under the law is less than the value of transferring a dollar to those who are eligible, in which case the appropriate value of a dollar of such transfer is  $\alpha\lambda$ , where  $\lambda < 1$ . Thus, the net social value impact of erroneous benefit payments is  $\$25,678,620 \times (\alpha\lambda - \beta)$ , which is positive if  $\lambda$  is close enough to 1, negative otherwise, and, finally, is never worse in social cost-benefit terms than  $-\$25,678,620 \times \beta$ . Our final cost-benefit formula is  $\$77,035,860 \times \alpha - \$94,965,106 \times \beta + \$25,678,620 \times (\alpha\lambda - \beta)$ , which, after some algebra, may be written as  $\$120,643,726 \times (\alpha - \beta) - \$43,607,866 \times \alpha + \$25,678,620 \times \alpha\lambda$ . If we assume that  $\alpha = 2$ ,  $\beta = 1.4$ , and  $\lambda = 0.5$  (so that a dollar of disability benefits paid to an ineligible person who is erroneously granted benefits on appeal has a social value of 50 cents), then the net social value is a gain of \$11 million. Raising the value of  $\lambda$  to 1 would yield a net social value that is roughly \$36 million. Reducing the value of  $\lambda$  to 0 instead yields a net social value that is a loss of roughly \$15 million. If we totally ignored the social costs related to judicial review—i.e., set  $\beta$  to 0—and assumed  $\lambda = 1$ , we would obtain a social value that is a gain of about \$205 million. If instead we kept the assumption of  $\beta = 1.4$  but totally ignored the social benefits—i.e., set  $\alpha$  to 0—we would obtain a social value that is a loss of about \$169 million. This discussion shows that even with relatively extreme assumptions about the parameter values necessary to measure the social costs and benefits of judicial review, the magnitude of the net social gain or loss would be somewhere in the neighborhood of \$100–\$200 million. That might sound like a lot of money, but it is a drop in the bucket in the context of the disability programs; SSDI alone accounted for \$147 billion in spending in 2015. SOC. SEC. ADMIN., 2016 OASDI TRUSTEES REPORT tbl.II.B1 (2016), [https://www.ssa.gov/OACT/TR/2016/II\\_B\\_cyoper.html#96807](https://www.ssa.gov/OACT/TR/2016/II_B_cyoper.html#96807) [<https://perma.cc/U264-HYTH>]. Thus, even our extreme assumptions yield net social gains or losses from judicial review of less than a tenth of a percent of the disability programs' overall spending. Our more reasonable assumptions yield estimates whose magnitudes are rounding error in the budgetary context.

claimants' lawyers when their clients prevail,<sup>83</sup> and the cost of federal judge time. In FY 2015, these resources funded a system of judicial review that corrected a maximum of about six-and-a-half errors per ALJ.<sup>84</sup> The SSA paid \$38,132,381.48 in EAJA fees in FY 2015.<sup>85</sup> This amount equals the salaries of about 240 ALJs, or 18% of their total number.<sup>86</sup> If spent on ALJs instead, this money alone could increase the ALJ corps by 18% and thereby enable the SSA to lower per capita case completion goals without increasing the backlog of undecided cases. If a lightened load led to even a modest improvement in decisional accuracy, i.e., seven fewer errors per ALJ,<sup>87</sup> then the resources spent on judicial review would yield fewer errors if redirected to error avoidance.<sup>88</sup>

This case for error avoidance rests on the assumption that the federal courts currently correct only a modest number of errors. If the number rises, the argument for an investment in error correction strengthens. In theory, Congress can control this number by resetting jurisdictional requirements and the federal courts' standard of review. It thereby could adjust the flow of cases to the federal courts. An endogeneity problem seems to exist. Whether Congress should increase the flow of cases to the federal courts depends on the value of the courts' error correction function. But the value of error correction depends on where Congress sets the dial to control the flow of cases to the federal courts. Also, very importantly, claimants' behavior might be different earlier in the process if there were no judicial review. For example, some claimants might not pursue appeals at earlier stages if they

---

83. On EAJA fees, see GELBACH & MARCUS, *supra* note 16, at 55.

84. In FY 2013, the country's 1,356 ALJs rendered 458,869 appealable decisions. SOC. SEC. ADMIN., ANNUAL STATISTICAL SUPPLEMENT TO THE SOCIAL SECURITY BULLETIN tbl.2.F8 (2015), <https://www.ssa.gov/policy/docs/statcomps/supplement/2014/supplement14.pdf> [<https://perma.cc/T7U5-JJE7>]; *Appeals to the AC as a Percentage of Appealable Hearing Level Dispositions*, SOC. SEC. ADMIN., [https://www.ssa.gov/appeals/DataSets/AC01\\_RR\\_Appealable\\_HO\\_Dispositions.html](https://www.ssa.gov/appeals/DataSets/AC01_RR_Appealable_HO_Dispositions.html) [<https://perma.cc/23T9-3XER>]. In FY 2015, the federal courts remanded 8,646 cases, or 6.4 remands per ALJ. *Court Remands as a Percentage of New Court Cases Filed*, SOC. SEC. ADMIN., [https://www.ssa.gov/appeals/DataSets/AC05\\_Court\\_Remands\\_NCC\\_Filed.html](https://www.ssa.gov/appeals/DataSets/AC05_Court_Remands_NCC_Filed.html) [<https://perma.cc/B9FD-EM46>]. For the purposes of this calculation, we assume that an ALJ decision issued in 2013 will get reviewed, if at all, by a federal judge in 2015.

85. *Social Security Administration Data for Equal Access to Justice Act Payments*, SOC. SEC. ADMIN., <https://www.ssa.gov/open/data/EAJA.xlsx> [<https://perma.cc/R4UM-GMV6>].

86. *Administrative Law Judge—Federal Salaries of 2016*, FEDERALPAY.ORG, <http://www.federalpay.org/employees/occupations/administrative-law-judge> [<https://perma.cc/WS9D-EE4L>].

87. On the relationship between quality and quantity, see GELBACH & MARCUS, *supra* note 16, at 72–73.

88. The math comes out the same way for immigration adjudication. In 2015, the federal courts of appeals decided 250 cases in favor of immigrants—about one per IJ. Guendelsberger, *supra* note 70, at 6. For information on the number of IJs during 2015, see Joshua Breisblatt, *Despite Immigration Judge Hiring, Court Backlogs Continue to Grow*, AM. IMMIGR. COUNCIL (July 27, 2016), <http://immigrationimpact.com/2016/07/27/despite-immigration-judge-hiring-court-backlogs-continue-grow/> [<https://perma.cc/3HGP-GLWT>].

knew there was no possible appeal to the federal courts, and they might have greater difficulty obtaining legal representation.

Other determinants of the bang for each buck invested in error correction, however, are exogenous. They depend on immutable institutional factors that constrain the federal courts' overall capacity to review agency decisions. Even under conditions that should prompt the most appeals, the federal courts receive few relative to the agency's caseload. In 2002, for example, the U.S. Attorney General announced changes to BIA processes to expedite agency review of IJ decisions.<sup>89</sup> Many believe that these "streamlining" changes degraded the BIA's review considerably by reducing the scrutiny it afforded IJ decisions.<sup>90</sup> BIA remands plummeted,<sup>91</sup> even as IJ decisional quality earned scathing criticism.<sup>92</sup> Cases flooded the courts of appeals,<sup>93</sup> rising from 1,760 in 2001 to 12,349 in 2005.<sup>94</sup> But even at the surge's peak, only about 5% of IJ decisions produced a federal court appeal.<sup>95</sup>

Attorney incentives are one such institutional factor that limits the federal courts' docket, regardless of where Congress sets the dial. Immigration and social security lawyers prefer to litigate before agency adjudicators rather than the federal courts. Disability and immigration cases generate only modest fees, so social security and immigration specialists often must have high volume practices.<sup>96</sup> For most lawyers, a federal court appeal takes much more time than an appearance before an IJ or ALJ.<sup>97</sup> Immigration lawyers typically represent clients for a flat fee,<sup>98</sup> an arrangement that should steer them toward less time intensive work (i.e., litigating in immigration court) than more (writing an appellate brief). Lawyers who represent social security claimants likewise have a strong

89. John D. Ashcroft & Kris W. Kobach, *A More Perfect System: The 2002 Reforms of the Board of Immigration Appeals*, 58 DUKE L.J. 1991, 1994–96 (2009); *see also* 8 C.F.R. § 1003.1 (2017).

90. *E.g.*, Stacy Caplow, *After the Flood: The Legacy of the "Surge" of Federal Immigration Appeals*, 7 NW. J.L. & SOC. POL'Y 1, 4–6 (2012).

91. *See* Legomsky, *supra* note 53, at 1668–70 (explaining the decline in pro-immigrant decisions by BIA reviewers caused by the regulatory changes).

92. *E.g.*, Benslimane v. Gonzales, 430 F.3d 828, 829–30 (7th Cir. 2005).

93. Bert I. Huang, *Lightened Scrutiny*, 124 HARV. L. REV. 1109, 1122–23 (2011).

94. Caplow, *supra* note 90, at 2–3.

95. In FY 2003, IJs decided 250,823 matters. U.S. DEP'T OF JUSTICE, EXEC. OFFICE FOR IMMIGRATION REVIEW, FY 2006 STATISTICAL YEAR BOOK C4 (2007). In 2005, immigrants filed 12,349 appeals from BIA decisions. ADMIN. OFFICE OF THE U.S. COURTS, JUDICIAL BUSINESS 2005, at 15.

96. Ingrid V. Eagly, *Remote Adjudication in Immigration*, 109 NW. U. L. REV. 933, 985 (2015).

97. GELBACH & MARCUS, *supra* note 16, at 55 & n.321; *see also* John R.B. Palmer et al., *Why Are So Many People Challenging Board of Immigration Appeals Decisions in Federal Court? An Empirical Analysis of the Recent Surge in Petitions for Review*, 20 GEO. IMMIGR. L.J. 1, 88 (2005).

98. David Gialanella, *The Skinny on Flat Fees*, A.B.A. J., July 2008, at 26, 26.

economic incentive to prefer agency work.<sup>99</sup> True, poor quality agency adjudication in some hearing offices may deepen the pool of potentially good appeals and make court work more attractive to lawyers.<sup>100</sup> But as long as lawyers can earn more litigating before IJs or ALJs, the supply of lawyers available to litigate federal court appeals in the areas where agency decision-making suffers may be insufficiently elastic to pick up the slack.

Attrition is perhaps an even more powerful institutional barrier to federal court. The extended process of adjudication and review within the agency can cause even those with meritorious appeals to give up before they reach the federal courts. By the time she can file an appeal in federal court, a disability benefits claimant may well have already spent more than 1,000 days pursuing her claim.<sup>101</sup> Although the time can vary considerably, an immigrant's case can easily languish for more than 1,000 days before an IJ and the BIA complete their review.<sup>102</sup> Beyond the time involved, carrots or sticks available to the agency can incentivize claimants or immigrants to forego an appeal. Prolonged detention encourages immigrants to eschew appeals and accept removal, presumably to end the misery of incarceration.<sup>103</sup> The SSA allows a previously denied claimant to file a new disability claim based upon a worsening of her condition, but she must abandon any pending appeal to do so.<sup>104</sup>

Finally, the federal courts' limited capacity to decide appeals in a manner consistent with deliberative judicial practice may ultimately impose an upper limit on how many cases they attract. As filings increase, judicial processes may change to such an extent that they increasingly resemble the fast, truncated adjudication that ALJs and IJs provide.<sup>105</sup> The Second and

---

99. See GELBACH & MARCUS, *supra* note 16, at 55 n.321 (noting the greater time demands of federal court appeals).

100. See Palmer et al., *supra* note 97, at 87–88 (explaining immigration lawyers' loss of faith and sense of injustice in BIA proceedings has improved their view of federal court appeals).

101. In FY 2015, the average processing time for a claim's determination at the initial level was 114 days. Reconsideration took 113 days on average, and a claim languished for 480 days before an ALJ's decision. 2015–2017 SOC. SEC. ADMIN. ANN. PERFORMANCE REP. 26 (2017). Appeals Council review took on average 386 days. *Social Security Administration Appeals Council Requests for Review Average Processing Time*, SOC. SEC. ADMIN., [www.ssa.gov/open/data/Appeals-Council-Avg-Proc-Time.html](http://www.ssa.gov/open/data/Appeals-Council-Avg-Proc-Time.html) [<https://perma.cc/WUH7-PRG2>].

102. In 2015, the average case languished for 504 days before an IJ decided it. *Immigration Court Processing Time by Outcome*, TRAC IMMIGR., [http://trac.syr.edu/phptools/immigration/court\\_backlog/](http://trac.syr.edu/phptools/immigration/court_backlog/) [<https://perma.cc/3JWK-GYUC>]. We could not find 2015 data for the BIA. In 2012, the BIA took an average of 485 days to decide an appeal filed by a nondetained immigrant. U.S. DEP'T OF JUSTICE, OFFICE OF THE INSPECTOR GEN., I-2013-001, MANAGEMENT OF IMMIGRATION CASES AND APPEALS BY THE EXECUTIVE OFFICE FOR IMMIGRATION REVIEW 43 (2012).

103. See BANKS MILLER ET AL., IMMIGRATION JUDGES AND U.S. ASYLUM POLICY 131–32 (2015) (finding that detained immigrants “are much less likely to appeal”).

104. SSR 11-1p, 76 Fed. Reg. 45,309, 45,310 (July 28, 2011).

105. See GELBACH & MARCUS, *supra* note 16, at 121–22 (describing the similarities in decision-making structures between magistrate judges with high caseloads and agency adjudicators); see also Robert A. Katzmann, *The Legal Profession and the Unmet Needs of the*

Ninth Circuits bore the brunt of the surge in immigration appeals after the BIA streamlining changes.<sup>106</sup> Starting in 2002, the Ninth Circuit made aggressive use of a case screening process that ultimately routed sixty percent of immigration cases to staff attorneys for a quick workup, followed by a brief oral presentation of each case to a screening panel of judges.<sup>107</sup> These judges, who typically did not review any materials in advance, decided 100–150 cases based on these presentations over a 2–3 day period.<sup>108</sup> The rate at which immigrants prevailed appears to have fallen sharply between 2002 and 2006.<sup>109</sup> Perhaps this assembly-line character dissuaded some appeals, as lawyers came to identify less of a difference between agency and court adjudication and perceived that increasing caseloads prompt courts to defer more to the agency’s decisions.<sup>110</sup>

### B. *The Regulatory and Critical Functions*

The opportunity cost problem weakens the contribution that the error correction function can make to the case for judicial review. But if courts not only correct errors but also induce agency adjudicators to avoid more in the

---

*Immigrant Poor*, 21 GEO. J. LEGAL ETHICS 3, 6 (2008) (describing the Second Circuit’s “non-argument calendar” for asylum cases, designed to handle the swell of immigration appeals).

106. Huang, *supra* note 93, at 1123–24.

107. Anna O. Law, *The Ninth Circuit’s Internal Adjudicative Procedures and Their Effect on Pro Se and Asylum Appeals*, 25 GEO. IMMIGR. L.J. 647, 673 (2011).

108. *Id.* at 675; see also Michael Kagan et al., *Buying Time? False Assumptions About Abusive Appeals*, 63 CATH. U. L. REV. 679, 702–05 (2014) (describing the Ninth Circuit screening system as efficient but expressing concern with its “heavy reliance on staff attorneys rather than judges.”).

109. The Administrative Office of the U.S. Courts provides termination data on “administrative appeals” and does not isolate immigration cases more specifically. During the time period of the surge, however, almost all of the change in the number of administrative appeals came from changes to the number of immigration cases appealed. ADMIN. OFFICE OF THE U.S. COURTS, JUDICIAL BUSINESS 2004 12–13. The following table includes the percentage of administrative appeals the Ninth Circuit either reversed or remanded out of the total number of administrative appeal terminations:

Year	Reversal/Remand Rate
2002	11%
2003	9.3%
2004	6.1%
2005	6.6%
2006	1.3%

Data come from Tables B-5 of the *Judicial Business of the United States Courts* for the years 2002–2006. See *Judicial Business of the United States Courts*, ADMIN. OFFICE OF THE U.S. COURTS, <http://www.uscourts.gov/statistics-reports/analysis-reports/judicial-business-united-states-courts> [<https://perma.cc/7GFM-CEPE>].

110. See Huang, *supra* note 93, at 1111–12 (discussing the concern, as Judge John Gibbons put it, that federal appellate 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”).

first place, then its claim to cost effectiveness strengthens. The Mashaw group suggested several ways by which court remands might play such a regulative or critical function. Judicial review might have an *in terrorem* effect on agency adjudication;<sup>111</sup> adjudicators might follow rules courts fashion, or an agency might use information gleaned from court remands to improve. As before, however, institutional determinants interfere with each of these possibilities.

*1. The In Terrorem Effect.*—An ALJ or an IJ focused on numbers alone has almost no reason to change her approach to decision-making just because a federal judge might reverse her. Only 2%–3% of ALJ decisions denying benefits produce a federal court remand.<sup>112</sup> The rate for IJs is even lower.<sup>113</sup> Another way to put it is to recall that federal courts remand roughly six cases per ALJ per year, whereas ALJs adjudicate about 540 claims per year.<sup>114</sup> Of course, agency adjudicators may vest outsized stock in the federal courts' opinions of their work. When ALJs sued to challenge the expectation that each decide 500–700 cases per year as a threat to their decisional independence, for example, their complaint alleged that the slipshod work this case completion goal required “injured” them because it “demeaned” them “in the eyes of the federal judiciary.”<sup>115</sup> To be taken seriously by Article III judges as black-robed colleagues might matter more to agency adjudicators than the odd remand here and there. Thus the threat of federal court review might alter their decision-making.<sup>116</sup>

But federal judge criticism may just as plausibly encourage indifference or hostility among agency adjudicators. For our report, we interviewed ALJs who work in a hearing office that generates few remands and ALJs from a hearing office that generates a lot of remands. The former reported much more positive views of federal court decision-making and commented on the

---

111. See Cass R. Sunstein, *On the Costs and Benefits of Aggressive Judicial Review of Agency Action*, 1989 DUKE L.J. 522, 527 (1989) (“Judicial review serves as a powerful ex ante deterrent to lawless or irrational agency behavior.”).

112. In FY 2013, ALJs issued 458,869 appealable decisions. SOC. SEC. ADMIN., *supra* note 84. In FY 2015, the federal courts remanded 8,646 cases. SOC. SEC. ADMIN., *supra* note 81.

113. In 2015, the federal courts of appeals reversed the BIA 250 times. See *supra* note 88. During FY 2013, immigration judges ordered removal in 95,838 removal proceedings. U.S. DEP'T OF JUSTICE, EXEC. OFFICE FOR IMMIGRATION REVIEW, FY 2015 STATISTICS YEARBOOK C2 (2016). This figure does not account for the IJ decisions in asylum cases. So the chances of an immigrant losing before an IJ, but eventually winning at a federal court of appeals, is less than 0.2%.

114. KRENT & MORRIS, *supra* note 6, at 6.

115. Complaint at 27, *Ass'n of Admin. Law Judges v. Colvin*, 2014 WL 789074 (N.D. Ill. 2014) (No. 13–cv–2925).

116. See Stuart L. Lustig et al., *Inside the Judges' Chambers: Narrative Responses from the National Association of Immigration Judges Stress and Burnout Survey*, 23 GEO. IMMIGR. L.J. 57, 71–72 (2008) (chronicling IJ complaints about threats to their esteem, including, for instance, the “[f]ear that every decision or proceeding may trigger a ‘personalized’ and scathing published criticism from the reviewing circuit court”).



instructional value of court remands; indeed, these ALJs prepare and circulate semi-annual memoranda summarizing all decisions from the district court to which most of their cases go.<sup>117</sup> In contrast, ALJs in the high-remand district complained that district judges have little understanding of or regard for agency processes and expressed no appreciation for district court feedback.<sup>118</sup> The hearing offices there lack any sort of structured process that would internalize learning from district court opinions.<sup>119</sup>

2. *The Precedential Effect.*—Any *in terrorem* effect or lack thereof is less significant if a court can impose precedent on the agency that forces it to improve. This “precedential effect” has long attracted criticism on grounds that generalist courts lack the requisite expertise and perspective to forge useful legal changes to a complex regulatory regime.<sup>120</sup> We take as a given the proposition that judges can craft wise opinions for these areas of law, a proposition that is necessary but not sufficient for the precedential effect to function. Regardless of this proposition, however, institutional features and incentives can render the actual effect of precedent on agency decision-making questionable for high volume adjudication.

First, reviewing courts might not have a lot of precedent-setting authority. This is clearly true when appeals first go to the federal district courts, whose decisions agencies can ignore as nonprecedential. It can also be true when courts of appeals review agency decisions, because an agency can narrow the range of issues for which the court can issue binding precedent. If an internal appellate tribunal issues an opinion that resolves an unsettled interpretive issue, as the full BIA does routinely,<sup>121</sup> courts must extend the decision deference if it meets certain criteria of authoritativeness.<sup>122</sup> An agency can control the lawmaking terrain even more completely by issuing legislative rules.<sup>123</sup>

---

117. GELBACH & MARCUS, *supra* note 16, at 119–20.

118. *Id.* at 120.

119. *Id.* at 121.

120. See Richard A. Posner, *Regulation (Agencies) Versus Litigation (Courts): An Analytical Framework*, in REGULATION VERSUS LITIGATION: PERSPECTIVES FROM ECONOMICS AND LAW 11, 19–20 (Daniel P. Kessler ed., 2011) (comparing courts to agencies and highlighting generalist courts’ lack of expertise and institutional restraint as disadvantages).

121. For all BIA precedent decisions, see *Attorney General and BIA Precedent Decisions*, U.S. DEP’T JUST., [www.justice.gov/eoir/ag-bia-decisions](http://www.justice.gov/eoir/ag-bia-decisions) [<https://perma.cc/4P6X-MNDB>].

122. *INS v. Aguirre-Aguirre*, 526 U.S. 415, 424 (1999); see also *United States v. Mead Corp.*, 533 U.S. 218, 226–27 (2001). *But see* *Mahn v. Attorney Gen.*, 767 F.3d 170, 173 (3d Cir. 2014) (holding that single-member BIA decisions do not warrant deference).

123. See, e.g., Revisions to the Rule Regarding the Evaluation of Medical Evidence, 82 Fed. Reg. 5844 (Jan. 18, 2017) (replacing the treating-physician rule by regulation); ADMIN. CONFERENCE OF THE U.S., SSA DISABILITY BENEFITS PROGRAMS: ASSESSING THE EFFICACY OF THE TREATING PHYSICIAN RULE 7–10 (2013) (describing Courts of Appeals’ creation of the treating-physician rules and later acknowledgment that the contours of the rule could be altered by the rulemaking process).

Second, agencies can resist control by judicial precedent when it does issue. Whether an agency can formally do so poses a complicated question, although the answer is probably no. To a greater or lesser degree, a number of agencies at one time or another have asserted a policy of “nonacquiescence,” whereby they reserve the right to treat appellate case law as nonbinding.<sup>124</sup> “Intercircuit nonacquiescence,” by which precedent binds adjudicators only within a circuit’s boundaries, is routine.<sup>125</sup> This practice necessarily weakens the power of judicial review to regulate agency behavior,<sup>126</sup> but no more than how circuit boundaries limit the force of any precedent. The Fourth Circuit cannot compel ALJs in Pasadena to follow its interpretation of the Social Security Act, but neither can the Fourth Circuit demand that police officers in Pasadena honor its understanding of the Fourth Amendment. The Supreme Court has never ruled on “intracircuit nonacquiescence,” the more problematic variant, whereby an agency denies that appellate precedent binds its decision-making even within that circuit’s boundaries. The lower federal courts have uniformly condemned the practice,<sup>127</sup> and neither the EOIR nor the SSA currently practices intracircuit nonacquiescence, at least formally.<sup>128</sup>

But acquiescence in judicial precedent does not necessarily happen automatically within an agency. The agency typically has a process to digest case law that, whether intentionally or unintentionally, can blunt the precedent’s force. When a court of appeals issues a published opinion that goes against the government in an immigration case, the EOIR’s Office of General Counsel must coordinate the agency’s response with the DOJ’s Office of Immigration Litigation and its own Office of the Chief Immigration Judge.<sup>129</sup> This “difficult” process<sup>130</sup> presumably can delay the opinion’s effect on IJ adjudication.

Something more than the unavoidable difficulty of bureaucratic coordination seems afoot in the SSA. Since 1985, the agency has required all ALJs within a circuit to follow that circuit’s precedent.<sup>131</sup> But ALJs do not

---

124. Samuel Estreicher & Richard L. Revesz, *Nonacquiescence by Federal Administrative Agencies*, 98 YALE L.J. 679, 692–718 (1989).

125. See, e.g., *Matter of Singh*, 25 I & N Dec. 670, 672 (B.I.A. 2012) (“We apply the law of the circuit in cases arising in that jurisdiction, but we are not bound by a decision of a court of appeals in a different circuit.”).

126. See ROBERT J. HUME, *HOW COURTS IMPACT FEDERAL ADMINISTRATIVE BEHAVIOR* 92–93 (2009) (explaining the impact of nonacquiescence on the force of precedent).

127. E.g., *Heartland Plymouth Court MI, LLC v. NLRB*, 838 F.3d 16, 25 (D.C. Cir. 2016); *Grant Med. Ctr. v. Burwell*, 204 F. Supp. 3d 68, 78–79 (D.D.C. 2016).

128. SSR 96-1p, 1996 WL 374182 (July 2, 1996); see *Peters v. Ashcroft*, 383 F.3d 302, 305 n.2 (5th Cir. 2004) (explaining who is bound by what precedent in immigration adjudication).

129. HUME, *supra* note 126, at 25.

130. *Id.*

131. MARTHA DERTHICK, *AGENCY UNDER STRESS: THE SOCIAL SECURITY ADMINISTRATION IN AMERICAN GOVERNMENT* 148 (1990).

simply read opinions on their own and decide whether and how a court has tweaked agency policy. The SSA instructs ALJs to ignore circuit decisions until the agency has determined that the decision conflicts with agency policy. Only then does the SSA issue an “acquiescence ruling” that directs ALJs to comply.<sup>132</sup> This threshold can cloak a de facto policy of intracircuit nonacquiescence. The agency can soft-pedal differences between precedent and its own policy, insisting that no conflict exists, and thereby instruct ALJs to ignore court decisions. In 2013, for example, the Fourth Circuit held that ALJs must give “substantial weight” to the Veterans Administration’s disability determination when a claimant with prior military service seeks social security benefits.<sup>133</sup> The social security ruling on the subject at the time was that the VA’s determination “cannot be ignored and must be considered,” an obligation that on its face seems weaker.<sup>134</sup> But the SSA never issued an acquiescence ruling for the Fourth Circuit’s opinion.<sup>135</sup> In fact, the agency has issued just over eighty acquiescence rulings during the acquiescence policy’s thirty-year history.<sup>136</sup> After an initial flurry of acquiescence rulings in the 1980s, when the policy began, the SSA’s pace has slowed markedly. Since 1990, the SSA has issued only three acquiescence rulings for the Second Circuit, for example,<sup>137</sup> and only three for the Seventh Circuit—a court that generated at least ten published opinions adverse to the agency in 2015 alone.<sup>138</sup>

The tactics agencies can use to limit case law’s significance matter less if agencies have no reason to resist regulation by precedent. But they do, for several reasons. First, agencies may believe that generalist courts inexpertly craft doctrine. Second, circuit-specific precedent can interfere with an agency’s effort to administer a single national policy uniformly across the country.<sup>139</sup> An agency may believe that justice lies in the consistent treatment

---

132. SSR 96-1p, 1996 WL 374182 (July 2, 1996).

133. *Bird v. Comm’r of Soc. Sec.*, 669 F.3d 337, 343 (4th Cir. 2012).

134. SSR 06-3p, 71 Fed. Reg. 45,593-03 (Aug. 9, 2006), *rescinded*, Revisions to Rules Regarding the Evaluation of Medical Evidence, 82 Fed. Reg. 5844, 5845 (Jan. 18, 2017).

135. In 2017, the SSA removed consideration of VA determinations entirely when it updated its medical evidence rules. Revisions to Rules Regarding the Evaluation of Medical Evidence, 82 Fed. Reg. at 5864, 5874 (codified at 20 C.F.R. §§ 404.1504, 416.904).

136. All acquiescence rulings are available here: *Acquiescence Rulings*, SOC. SEC. ADMIN., [www.ssa.gov/OP\\_Home/rulings/ar-toc.html](http://www.ssa.gov/OP_Home/rulings/ar-toc.html) [<https://perma.cc/2GA5-V4NS>].

137. *See Acquiescence Rulings: Second Circuit Court*, SOC. SEC. ADMIN., [www.ssa.gov/OP\\_Home/rulings/ar/02/AR02toc.html](http://www.ssa.gov/OP_Home/rulings/ar/02/AR02toc.html) [<https://perma.cc/G265-Z25Z>].

138. *See id.* The Seventh Circuit published the following cases adverse to the SSA: *Hill v. Colvin*, 807 F.3d 862, 863 (7th Cir. 2015); *Alaura v. Colvin*, 797 F.3d 503, 508 (7th Cir. 2015); *Stepp v. Colvin*, 795 F.3d 711, 713 (7th Cir. 2015); *Price v. Colvin*, 794 F.3d 836, 841 (7th Cir. 2015); *Varga v. Colvin*, 794 F.3d 809, 810 (7th Cir. 2015); *Engstrand v. Colvin*, 788 F.3d 655, 656 (7th Cir. 2015); *Voight v. Colvin*, 781 F.3d 871, 879–80 (7th Cir. 2015); *Adaire v. Colvin*, 778 F.3d 685, 688 (7th Cir. 2015); *Hall v. Colvin*, 778 F.3d 688, 691 (7th Cir. 2015); *Minnick v. Colvin*, 775 F.3d 929, 939 (7th Cir. 2015).

139. Harold H. Bruff, *Coordinating Judicial Review in Administrative Law*, 39 UCLA L. REV. 1193, 1205–06 (1992).

of regulated entities or beneficiaries, regardless of what courts say in different parts of the country.<sup>140</sup> Also, the administration of a policy that splinters into dozens of geographically determined variants, to be applied by hundreds of different adjudicators, could prove impossibly difficult to administer. ALJs and IJs have earned harsh criticism for decisional inconsistencies.<sup>141</sup> While IJ disparities remain stubborn and notorious,<sup>142</sup> the SSA has undertaken significant efforts to identify reasons for ALJ idiosyncrasy and to counteract them.<sup>143</sup> If the SSA instructed ALJs to abide by circuit and district precedent, the agency would invite ALJs to draw their own judgment about governing policy and complicate its efforts to get more than 1,000 adjudicators on roughly the same policy-compliant page. For this reason,<sup>144</sup> the SSA has instructed ALJs and decision writers “not to consider any district court decisions.”<sup>145</sup>

If an agency is recalcitrant, Congress can structure judicial review to maximize courts’ power to create a precedential effect. As some have proposed for social security disability claims litigation, Congress can require that appeals go directly to circuit courts, not district courts, and it can steer all appeals to a single circuit.<sup>146</sup> Doing so would undermine a key argument for nonacquiescence: that different instructions from geographically dispersed courts would flummox an agency’s effort to administer a single national policy. But this arrangement would require either significantly less litigation, a dramatic change to judicial standards for acceptable decision-making, or a huge increase in the size of the designated appellate court. When Congress contemplated legislation to send all immigration appeals to the Federal Circuit, the Federal Circuit’s chief judge estimated that judicial time for decision-making would plummet to an hour-and-a-half per case as a result.<sup>147</sup> Were Congress to centralize all of the disability appeals currently pending before the regional circuits in the Federal Circuit, its caseload would spike by 25%, assuming no changes in claimant behavior; if all disability cases pending before the district courts went to the Federal Circuit, the latter

---

140. GELBACH & MARCUS, *supra* note 16, at 74–75.

141. *Id.* at 84–85 & n.444.

142. *E.g.*, U.S. GOV’T ACCOUNTABILITY OFF., GAO-17-72, ASYLUM: VARIATIONS EXIST IN OUTCOMES OF APPLICATIONS ACROSS IMMIGRATION COURTS AND JUDGES 2, 17 (2016).

143. Gerald K. Ray & Jeffrey S. Lubbers, *A Government Success Story: How Data Analysis by the Social Security Appeals Council (with a Push from the Administrative Conference of the United States) Is Transforming Social Security Disability Adjudication*, 83 GEO. WASH. L. REV. 1575, 1606 (2015).

144. Memorandum from Debra Bice, Chief Admin. Law Judge, to All Administrative Law Judges and All Senior Attorneys 2 (Jan. 11, 2013) (on file with authors); *see also* GELBACH & MARCUS, *supra* note 16, at 76 (quoting the Bice memorandum).

145. Bice, *supra* note 144, at 1–2.

146. *E.g.*, Richard E. Levy, *Social Security Disability Determinations: Recommendations for Reform*, 1990 BYU L. REV. 461, 512–17 (1990).

147. *Immigration Litigation Reduction: Hearing Before the S. Comm. on the Judiciary*, 109th Cong. 4 (2006) (statement of Paul R. Michel, C.J.).

would have to grow by dozens of judges to keep its caseload at manageable levels.<sup>148</sup>

3. *Feedback.*—Whether binding or not, court decisions can serve as a valuable source of feedback and thereby discharge a critical function. An agency can always examine its wins and losses in court to look for ways to improve. But several institutional contrasts between courts and agencies may reduce agency incentives to do so.

One involves institutional goals. On a superficial level, agencies and courts share the same goal: the accurate and efficient implementation of the relevant regulatory regime. On another, however, these goals diverge. Agencies attempt to meet standards for decisional quality, but quantity—case completion goals, production quotas, and so forth—matter just as much, if not more, in measures of agency performance.<sup>149</sup> Quality conflicts with quantity, for obvious reasons.<sup>150</sup> ALJs surely could generate better decisions with half as many claims to adjudicate, but claimants would then wait twice as long for a hearing. The SSA is legitimately concerned with the injustice of

---

148. In FY 2014, the Court of Appeals for Veterans Claims decided about 175 petitions and appeals on the merits per “active judge.” 2014 U.S. CT. APP. VETERAN CLAIMS ANN. REP. 5, [www.uscourts.cavc.gov/documents/FY2014AnnualReport06MAR15FINAL.pdf](http://www.uscourts.cavc.gov/documents/FY2014AnnualReport06MAR15FINAL.pdf) [https://perma.cc/Y6ZD-R8GP]. A court of appeals handling all 20,000 social security cases presently filed in the district courts would have to have 114 judges dedicated just to this litigation to have an equivalent caseload. The CAVC has attracted criticism for its backlog. *E.g.*, Jerry Markon, *Veterans Court Faces Backlog that Continues to Grow*, WASH. POST (Apr. 22, 2011), [https://www.washingtonpost.com/politics/veterans-court-faces-backlog-that-continues-to-grow/2011/04/15/AFFaavRE\\_story.html?utm\\_term=.df2fbcc9128](https://www.washingtonpost.com/politics/veterans-court-faces-backlog-that-continues-to-grow/2011/04/15/AFFaavRE_story.html?utm_term=.df2fbcc9128) [https://perma.cc/99B5-44BS].

149. See Eric Biber, *Too Many Things to Do: How to Deal with the Dysfunctions of Multiple-Goal Agencies*, 33 HARV. ENV'T L.L. REV. 1, 12 (2009) (arguing that agencies “tend to overproduce on the goals that are complements and the goals that are easily measured”); STAFF OF H. COMM. ON OVERSIGHT & GOV'T REFORM, 113TH CONG., MISPLACED PRIORITIES: HOW THE SOCIAL SECURITY ADMINISTRATION SACRIFICED QUALITY FOR QUANTITY IN THE DISABILITY DETERMINATION PROCESS 49 (2014) (asserting that “many of [the agency’s] failings are attributable to the agency’s development of a factory-like production process that ignores the quality of ALJ decisions”).

150. *Ass’n of Admin. Law Judges v. Colvin*, 777 F.3d 402, 404–05 (7th Cir. 2015); Stephen W. Gilliland & Ronald S. Landis, *Quality and Quantity Goals in a Complex Decision Task: Strategies and Outcomes*, 77 J. APPLIED PSYCHOL. 672, 680 (1992).

a claim's being unreasonably delayed.<sup>151</sup> It faces constant and enduring scrutiny for its claims backlog,<sup>152</sup> as does the EOIR.<sup>153</sup>

Agencies have the complex task of successfully managing the tradeoff between quantity and quality. Typically, the federal courts do not shoulder the same obligation to generate large numbers of decisions quickly.<sup>154</sup> Agencies constantly monitor adjudicator productivity and evaluate performance in terms of it.<sup>155</sup> The institutional culture of the federal judiciary would not permit the same sort of pressure on individual judges.<sup>156</sup> Moreover, the federal courts do not endure the same legislative and public scrutiny for their pace of decision-making that agencies routinely confront. Federal judges can therefore render particularized justice tailored to the circumstances of an individual case without significant regard for production quotas. Differences in resources available to decide cases exacerbate the significance of these contrasting goals. An ALJ deciding up to fifty cases per month has a fundamentally different job than a federal judge and her clerk, who can deliberate on a case for a week.<sup>157</sup>

An agency adjudicator might treat judicial feedback as unhelpful if it does not account for her need to produce decisions quickly under severe resource constraints. An example involves the enforcement of subpoenas ALJs issue to medical providers for relevant records.<sup>158</sup> To some federal

---

151. See *Nash v. Bowen*, 869 F.2d 675, 680 (2d Cir. 1989) (discussing “the Secretary’s policy of setting a minimum number of dispositions an ALJ must decide in a month” and agreeing “with the district court that reasonable efforts to increase the production levels of ALJs are not an infringement of decisional independence”).

152. E.g., David A. Fahrenthold, *At Social Security Office with a Million-Person Backlog, There’s a New Chief*, WASH. POST (July 23, 2015), [https://www.washingtonpost.com/news/federal-eye/wp/2015/07/23/at-social-security-office-with-a-million-person-backlog-theres-a-new-chief/?utm\\_term=.bf7210f69698](https://www.washingtonpost.com/news/federal-eye/wp/2015/07/23/at-social-security-office-with-a-million-person-backlog-theres-a-new-chief/?utm_term=.bf7210f69698) [<https://perma.cc/Y7QK-2LND>].

153. E.g., Julia Preston, *Deluged Immigration Courts, Where Cases Stall for Years, Begin to Buckle*, N.Y. TIMES (Dec. 1, 2016), <https://www.nytimes.com/2016/12/01/us/deluged-immigration-courts-where-cases-stall-for-years-begin-to-buckle.html> [<https://perma.cc/RPC8-P3XW>].

154. See RICHARD A. POSNER, *HOW JUDGES THINK* 140–41 (2008) (commenting on district judges’ sensitivity to delays in deciding motions but noting that there is no sanction for delays). But see Diarmuid F. O’Scainnlain, *Striking a Devil’s Bargain: The Federal Courts and Expanding Caseloads in the Twenty-First Century*, 13 LEWIS & CLARK L. REV. 473, 476 (2009) (observing that about one-third of Ninth Circuit cases get decided by a “screening panel” of judges that spend four to nine minutes on each after a workup by a staff attorney).

155. E.g., GELBACH & MARCUS, *supra* note 16, at 41 (describing an online tracking program called “How MI Doing” that allows an ALJ to see the number of cases she has decided and her remand rate).

156. See Richard J. Pierce, Jr., *Political Control Versus Impermissible Bias in Agency Decisionmaking: Lessons from Chevron and Mistretta*, 57 U. CHI. L. REV. 481, 516–17 (1990) (commenting on federal judges’ aversion, based partly on their “self-image,” to interference with their decision-making processes).

157. GELBACH & MARCUS, *supra* note 16, at 74.

158. See 20 C.F.R. §§ 404.950(d)(1), 416.1450(d)(1) (2017); SOC. SEC. ADMIN., HEARINGS, APPEALS, AND LITIGATION LAW MANUAL I-2-5-78 [hereinafter HALLEX], [https://www.ssa.gov/OP\\_Home/hallex/I-02/I-2-5-78.html](https://www.ssa.gov/OP_Home/hallex/I-02/I-2-5-78.html) [<https://perma.cc/XF4A-GJML>].

courts, especially in pro se cases, the mere issuance of a subpoena does not discharge the ALJ's obligation to "develop" the record<sup>159</sup> when the person or entity being subpoenaed does not respond.<sup>160</sup> An ALJ who seeks a subpoena's enforcement, however, must trigger a cumbersome, time-intensive process.<sup>161</sup> The SSA may follow through on a particular court remand requiring a subpoena's enforcement. But the agency is not likely to act on this feedback more generally and institutionalize a subpoena enforcement policy, given the demands of its caseload.<sup>162</sup>

A second institutional difference might affect the filter through which adjudicators view court feedback, countering its potency. Agency adjudicators might feel obliged to honor aggregate-level, agency-wide policy goals that courts do not countenance.<sup>163</sup> A need to "protect the fund" and the overall health of the social security program might influence ALJ decision-making in individual cases.<sup>164</sup> Observers have long commented on the uncomfortable placement of IJs within the DOJ, suggesting that this institutional arrangement may skew decision-making in favor of strict enforcement.<sup>165</sup> Federal judges face no such aggregate-level pressure for the successful administration of a complex regulatory regime.

Two other institutional differences can also undermine guidance derived from judicial opinions. The first is the baseline problem described above. A

159. 20 C.F.R. §§ 404.1512(d)–(e), 416.912(d)–(e).

160. *E.g.*, *Brandow v. Comm'r of Soc. Sec.*, No. 1:05–CV–09171 NPMVEB, 2009 WL 2971543, at \*5 n.6 (N.D.N.Y. Sept. 11, 2009); *Suriel v. Comm'r of Soc. Sec.*, Civ. No. 05-1218, 2006 WL 2516429, at \*6 (E.D.N.Y. Aug. 29, 2006); *Sanchez v. Barnhart*, 329 F. Supp. 2d 445, 450–51 (S.D.N.Y. 2004). *But see* *Friedman v. Astrue*, No. 07 Civ. 3651, 2008 WL 3861211, at \*8–9 (S.D.N.Y. Aug. 19, 2008) (abiding by a Second Circuit holding that an "ALJ's decision to enforce a subpoena on an unresponsive party is discretionary"); *Serrano v. Barnhart*, No. 02 Civ. 6372, 2005 WL 3018256, at \*3–4 (S.D.N.Y. Nov. 1, 2005) (finding that a mandate to enforce a subpoena would be a "tremendous and undue burden" on an ALJ).

161. HALLEX, *supra* note 158, at I-2-5-82, [https://www.ssa.gov/OP\\_Home/hallex/I-02/I-2-5-82.html](https://www.ssa.gov/OP_Home/hallex/I-02/I-2-5-82.html) [<https://perma.cc/4N4E-ZRW3>]; *see also* *Yancey v. Apfel*, 145 F.3d 106, 113 (2d Cir. 1998) (expressing concerns over "the financial and administrative burdens of processing disability claims" that a rule requiring the SSA to subpoena treating physicians at the claimant's behest would entail).

162. *See* GELBACH & MARCUS, *supra* note 16, at 19–21 (juxtaposing the difficulties of the subpoena process with ALJs' institutional "just in time" approach towards case review).

163. Jerry L. Mashaw, *Organizing Adjudication: Reflections on the Prospect for Artisans in the Age of Robots*, 39 UCLA L. REV. 1055, 1056 (1992).

164. GELBACH & MARCUS, *supra* note 16, at 78–79; *see also* D. Randall Frye, *Statement of the Association of Administrative Law Judges, Committee on Ways and Means, Subcommittee on Social Security*, 33 J. NAT'L ASS'N ADMIN. L. JUDICIARY 35, 42–43 (2013) (June 27, 2012) (stating that "having the judge defend the Trust Fund as well as the claimant's interest . . . places the judge in an untenable situation").

165. *E.g.*, U.S. COMM'N ON IMMIGRATION REFORM, *BECOMING AN AMERICAN: IMMIGRATION AND IMMIGRANT POLICY* 178 (1997) (advocating that citizenship and immigration adjudications be moved from the DOJ to the State Department); AM. BAR ASS'N HOUSE OF DELEGATES, *RESOLUTION 114F*, at 4–5 (2010) (describing criticism of the placement of immigration adjudication under the purview of the DOJ and advocating for fundamental "restructuring").

nitpicky remand of a clearly meritless claim might lead the ALJ to discount the district court's order, and perhaps future ones, as uninformed. Second, the agency might explicitly discourage its adjudicators from considering court remands as a source of feedback, concerned that doing so might create discrepancies in adjudicators' understandings of policy-compliant decision-making.

Whatever the reason, the SSA presently does little as an agency<sup>166</sup> to mine district court remand decisions for instruction. An ALJ who gets remanded will see the decision, but the decision writer who drafted it will not.<sup>167</sup> Neither does the Appeals Council analyst nor the appellate adjudicator. The EOIR has no mechanism in place to ensure that staff attorneys involved in a decision that gets remanded see the court opinion and learn from it.<sup>168</sup>

\* \* \*

The foregoing dwells on the many institutional impediments that interfere with judicial review's corrective, regulative, and critical functions. The story may not be quite so bleak. In particular instances courts may discharge one or more of the functions more successfully. The body of immigration law that IJs administer, for instance, owes a good deal to federal circuit precedent. Also, the case for judicial review should not depend upon the justificatory force of any single function in isolation but rather the cumulative contributions that courts can make. Courts may not correct errors more efficiently than adjudicators can avoid them, but if they can rectify some mistakes and exert some regulative influence, however limited, then perhaps the case for judicial review of high volume agency adjudication strengthens.

Those who have studied high volume agency adjudication most closely remain unconvinced. The Mashaw group favored the replacement of Article III review of ALJ decision-making with a specialized social security court,<sup>169</sup> a recommendation seconded by distinguished commentators.<sup>170</sup> When Congress caved to judicial pressure and created judicial review for veterans benefits adjudication in 1988, it opted for a specialized Article I court.<sup>171</sup> A proposal to jettison review of IJ decisions by the regional courts

---

166. ALJs on their own in some instances have created organized methods of deriving feedback from district court decisions. GELBACH & MARCUS, *supra* note 16, at 119–20.

167. *Id.* at 174–75.

168. Lenni B. Benson, *You Can't Get There From Here: Managing Judicial Review of Immigration Cases*, 2007 U. CHI. LEGAL F. 405, 427 (2007).

169. MASHAW ET AL., *supra* note 15, at 146–50.

170. Verkuil & Lubbers, *supra* note 15, at 776, 781–82.

171. Michael Serota & Michelle Singer, *Veterans' Benefits and Due Process*, 90 NEB. L. REV. 388, 396 (2011).



of appeals gained traction in Congress in the mid-2000s.<sup>172</sup> Many clearly continue to believe that whatever benefits Article III review brings to high volume agency adjudication, they fall short of justifying it.

### III. Problem-Oriented Oversight

Judgments about judicial review's wisdom are incomplete because existing accounts of its role supervising high volume agency adjudication have overlooked a key function courts can perform. This function has something important to do with an interesting dynamic apparent in the case law this litigation generates. Often boring and repetitive, appeals typically yield cookie-cutter opinions of little significance.<sup>173</sup> Not infrequently, however, judges break this tedium with extraordinary commentary on patterns or trends they have observed. Identifying a set of ALJ decisions he found troubling, for example, a magistrate judge recently described some social security proceedings as "border[ing] on madness."<sup>174</sup> In a separate opinion released the same day, he denounced ALJ decisions as "littered with recurring issues" and lampooned social security appeals as "Groundhog Day."<sup>175</sup> Perhaps such statements, which are legion in immigration opinions<sup>176</sup> and not uncommon in social security cases,<sup>177</sup> are little more than outbursts of judicial frustration. But Article III judges tend to keep their powder pretty dry, so we interpret this sort of commentary as purposeful.

From time to time, judges try to influence agency decision-making through means beyond the correction of discrete errors in individual cases or the issuance of binding precedent. A comparison provides some insight into what courts might be up to. Congress has a lot of tools at its disposal to influence agency behavior.<sup>178</sup> One important one is a form of oversight, by which legislators assemble information on an agency, then comment publicly

---

172. *E.g.*, *Immigration Litigation Reduction: Hearing Before the S. Comm. on the Judiciary*, 109th Cong. 2 (2006) (statement of Sen. Arlen Specter, Chairman, S. Comm. on the Judiciary).

173. *E.g.*, MASHAW ET AL., *supra* note 15, at 140.

174. *Wallace v. Colvin*, 193 F. Supp. 3d 939, 941 (N.D. Ill. 2016).

175. *Booth v. Colvin*, No. 14 CV 50347, 2016 WL 3476700, at \*1 (N.D. Ill. June 27, 2016).

176. *See Legomsky*, *supra* note 53, at 1645 (referring to "the unprecedented scathing criticisms that so many U.S. courts of appeals have leveled at EOIR").

177. *E.g.*, *Hughes v. Astrue*, 705 F.3d 276, 279 (7th Cir. 2013) (admonishing the SSA and DOJ to "do better than they did in this case"); *Hardman v. Barnhart*, 362 F.3d 676, 679 (10th Cir. 2004) (reiterating the demand for more than standard boilerplate language because it "fails to inform [the court] in a meaningful, reviewable way of the specific evidence the ALJ considered in determining that claimant's complaints were not credible"); *Batista v. Colvin*, Civ. No. 13-4185, 2014 U.S. Dist. LEXIS 80576, at \*6 (E.D.N.Y. June 11, 2014) (demanding that "[j]ust once, [the court] would like to see an ALJ write" specific reasons for rejecting a plaintiff's credibility); *Freismuth v. Astrue*, 920 F. Supp. 2d 943, 954 (E.D. Wis. 2013) (describing the SSA Commissioner's brief that didn't include a citation to case law as "insulting" to the court because the "Commissioner's counsel can neither appropriately screen [ALJ decisions] nor adequately brief them").

178. Jack M. Beerman, *Congressional Administration*, 43 SAN DIEGO L. REV. 61, 69-70 (2006).

and critically on its performance.<sup>179</sup> Although in theory backed by the threat of a budget cut or some other legislative sanction, these congressional interventions can derive force simply from the informal pressure they generate.<sup>180</sup> We argue that courts attempt something similar, what we call “problem-oriented oversight,” when they decide certain appeals.

Courts engage in problem-oriented oversight when they identify and respond to “problems,” defined either as flawed administrations of policy by the agency, or as the agency’s nonresponse to an entrenched decision-making pathology. This Part distinguishes problem-oriented oversight from existing models of agency oversight and explains how courts engage in the task. Part IV examines the institutional factors that determine whether this function can succeed.

#### A. *Models of Agency Oversight*

The notion that judicial review functions as a type of agency oversight is hardly novel.<sup>181</sup> What exactly this oversight is and how courts conduct it in the context of high volume agency adjudication, however, have attracted little examination.

We begin with what Mariano-Florentino Cuellar aptly calls an “incredibly durable framework for thinking about legislative oversight of the bureaucracy,”<sup>182</sup> a subject that has garnered more study than court-based oversight. This canonical framework describes oversight in terms of two models.<sup>183</sup> When Congress engages in “police patrol oversight,” it surveys a large number of agency decisions or actions, selected at random, to determine if the agency is functioning properly.<sup>184</sup> Like a police officer cruising a neighborhood, this oversight happens when, “at its own initiative, Congress examines a sample of executive-agency activities, with the aim of detecting and remedying any violations of legislative goals and, by its surveillance, discouraging such violations.”<sup>185</sup> Police patrol oversight is proactive and often regular and ongoing.<sup>186</sup> Examples include making an agency submit annual reports to Congress, obliging agency officials to appear at committee

---

179. *Id.* at 122–23, 125.

180. Douglas Kriner, *Can Enhanced Oversight Repair “the Broken Branch”?*, 89 B.U. L. REV. 765, 784–85 (2009).

181. See, e.g., David S. Law, *A Theory of Judicial Power and Judicial Review*, 97 GEO. L.J. 723, 747–78 (2009) (arguing courts make efficient and effective monitors of government conduct).

182. Mariano-Florentino Cuellar, *Auditing Executive Discretion*, 82 NOTRE DAME L. REV. 227, 297 (2006).

183. Matthew McCubbins & Thomas Schwartz, *Congressional Oversight Overlooked: Police Patrols Versus Fire Alarms*, 28 AM. J. POL. SCI. 165, 165–66 (1984).

184. *Id.* at 166.

185. *Id.*

186. LINDA L. FOWLER, *WATCHDOGS ON THE HILL: THE DECLINE OF CONGRESSIONAL OVERSIGHT OF U.S. FOREIGN RELATIONS* 134–35, 139 (2015).

hearings in connection with an annual budget request, and submitting an agency to examination by the Government Accountability Office.<sup>187</sup>

“Fire alarm oversight,” the second model, responds to institutional constraints, including high costs and inconstant legislator attention, that in theory limit the efficacy of police patrols.<sup>188</sup> Rather than itself gather and sift through large amounts of information about agency performance to find possible problems, “Congress establishes a system of rules, procedures, and informal practices that enable individual citizens and organized interest groups to examine administrative decisions . . . , to charge executive agencies with violating congressional goals, and to seek remedies from agencies, courts, and Congress itself.”<sup>189</sup> Such mechanisms are “fire alarms” that third parties can ring and thereby direct oversight attention to agency misconduct. Thus, this oversight is episodic and reactive.

A recent disability benefits scandal nicely illustrates fire alarm oversight. David Daugherty, an ALJ in Huntington, West Virginia, granted benefits in 1,280 of the 1,284 cases he decided in FY 2010.<sup>190</sup> This was the sixth year in a row in which Daugherty had decided more than 1,000 cases;<sup>191</sup> it came amidst a stunning growth in the nation’s disability rolls, and in a year when ALJs granted benefits in more than 70% of cases they decided on the merits.<sup>192</sup> Protected by a statutory safe harbor,<sup>193</sup> a prototypical fire alarm,<sup>194</sup> a whistleblower contacted the *Wall Street Journal* to bring Daugherty’s practice of rubber-stamping disability benefits claims to light.<sup>195</sup> The article

---

187. Beerman, *supra* note 165, at 66–67.

188. McCubbins & Schwartz, *supra* note 183, at 168.

189. *Id.* at 166.

190. Damian Paletta, *Disability-Claim Judge Has Trouble Saying “No”*, WALL STREET J. (May 19, 2011), <https://www.wsj.com/articles/SB10001424052748704681904576319163605918524> [<https://perma.cc/D5XU-4KM2>].

191. STAFF OF S. COMM. ON HOMELAND SEC. & GOVERNMENTAL AFFAIRS, 113TH CONG., HOW SOME LEGAL, MEDICAL, AND JUDICIAL PROFESSIONALS ABUSED SOCIAL SECURITY DISABILITY PROGRAMS FOR THE COUNTRY’S MOST VULNERABLE: A CASE STUDY OF THE CONN LAW FIRM 34 (2013).

192. SOC. SEC. ADVISORY BD., ASPECTS OF DISABILITY DECISION MAKING: DATA AND MATERIALS 12 (2012); CTR. ON BUDGET & POLICY PRIORITIES, CHART BOOK: SOCIAL SECURITY DISABILITY INSURANCE 6–7 (2017), <https://www.cbpp.org/sites/default/files/atoms/files/7-21-14socsec-chartbook.pdf> [<https://perma.cc/8C8B-T7SK>].

193. 5 U.S.C. § 2302(b)(8) (2012).

194. David E. Pozen, *Freedom of Information Beyond the Freedom of Information Act*, 165 U. PA. L. REV. 1097, 1143 (2017).

195. Paletta, *supra* note 190; *see also* Press Release, Soc. Sec. Admin., Office of the Inspector Gen., Former SSA Chief Administrative Law Judge Pleads Guilty to Conspiracy to Retaliate Against Informant (June 13, 2016), <https://oig.ssa.gov/audits-and-investigations/investigations/june13-andrus-guilty-plea> [<https://perma.cc/PAY6-VLVR>] (describing the prosecution of the former chief ALJ in the Huntington office for conspiring to retaliate against an SSA employee who was an informant for federal investigators).

prompted several congressional hearings<sup>196</sup> and at least two committee reports.<sup>197</sup> What emerged was criticism that the SSA, focused on case-completion goals above all else, turned a blind eye to ALJs “paying down” a huge backlog of claims.<sup>198</sup> Daugherty eventually pleaded guilty to felony charges, admitting that he took kickbacks from a local social security lawyer who received fees when Daugherty granted his clients’ claims.<sup>199</sup> Although the SSA denied the blind-eye charge, it made significant changes, at least partially in response to congressional scrutiny.<sup>200</sup> The ALJ claim allowance rate declined sharply, to 48%, by 2013.<sup>201</sup>

Although developed to describe versions of Congressional oversight, the police patrol and fire alarm models have come to serve as descriptions of how a range of overseers, including courts, can supervise agencies.<sup>202</sup> Judicial review has traditionally been treated as a component in a fire alarm system, with courts either as the oversight institution itself, or with courts serving as a forum where aggrieved third parties can ring a fire alarm and thereby trigger oversight.<sup>203</sup>

---

196. See, e.g., *Social Security Administration Oversight: Examining the Integrity of the Disability Determination Appeals Process: Hearing Before the H. Committee on Oversight and Government Reform*, 113th Cong. 29 (2014) (statement of Lamar Smith, Chairman, H. Comm. On the Judiciary) (indicating Congress first learned of the matter from the *Wall Street Journal*). See generally *Social Security Disability Benefits: Did a Group of Judges, Doctors, and Lawyers Abuse Programs for the Country's Most Vulnerable?: Hearing Before the S. Comm. on Homeland Sec. & Governmental Affairs*, 113th Cong. (2013) [hereinafter *Hearing on SSDI Abuse*] (repeatedly discussing the *Wall Street Journal* article).

197. STAFF OF H. COMM. ON OVERSIGHT & GOV'T REFORM, *supra* note 149, at 6 & n.6; STAFF OF S. COMM. ON HOMELAND SEC. & GOVERNMENTAL AFFAIRS, *supra* note 191, at 6.

198. See STAFF OF H. COMM. ON OVERSIGHT & GOV'T REFORM, *supra* note 149, at 5–7 (identifying the SSA's emphasis on quantity over quality in ALJ decisions); Stephen Olemacher, *Judges Tell Lawmakers They Are Urged to Approve Social Security Disability Claims*, WASH. POST (June 27, 2013), [https://www.washingtonpost.com/politics/judges-tell-lawmakers-they-are-urged-to-approve-social-security-disability-claims/2013/06/27/ea990a7e-df66-11e2-b2d4-ea6d8f477a01\\_story.html?utm\\_term=.3bccda0a508a](https://www.washingtonpost.com/politics/judges-tell-lawmakers-they-are-urged-to-approve-social-security-disability-claims/2013/06/27/ea990a7e-df66-11e2-b2d4-ea6d8f477a01_story.html?utm_term=.3bccda0a508a) [<https://perma.cc/7VUA-C7QG>] (reporting ALJs describing a system where judges were urged to grant claims for the sake of reducing the case backlog).

199. Stephen Dinan, *Judge Pleads Guilty in Massive Social Security Fraud Case*, WASH. TIMES (May 14, 2017), <http://www.washingtontimes.com/news/2017/may/14/david-b-daugherty-pleads-guilty-in-massive-social-/> [<https://perma.cc/H8AR-77QH>].

200. *Social Security Disability Benefits: Did a Group of Judges, Doctors, and Lawyers Abuse Programs for the Country's Most Vulnerable?: Hearing Before the S. Comm. on Homeland Sec. & Governmental Affairs*, 113th Cong. 127 (2013) (statement of Debra Bice, Chief ALJ, Social Security Administration).

201. 2017 SOC. SEC. ADMIN. BUDGET JUSTIFICATION, *supra* note 74, at 144.

202. Jacob E. Gersen & Anne Joseph O'Connell, *Hiding in Plain Sight? Timing and Transparency in the Administrative State*, 76 U. CHI. L. REV. 1157, 1172 (2009).

203. *Id.*; Law, *supra* note 181, at 747–48.

*B. The Limits of the Fire Alarm Model*

When Richard Posner castigated the “Immigration Court” as “the least competent federal agency” in a 2016 opinion,<sup>204</sup> perhaps he meant his harsh words as an attempt at fire alarm oversight. A third party, the immigrant facing removal, brought an alleged agency problem to a court and got Judge Posner to respond vociferously. But for several reasons the fire alarm model imperfectly describes what courts do. First, courts review large numbers of cases, most of which were either acceptably decided or at worst marred by random error. Fire alarm oversight is premised on the notion that third parties screen agency decisions for the overseer, finding agency flaws for a court, or a legislature motivated by a court, to fix. If this is so, the mechanism would seem to fit high volume agency adjudication poorly. Indeed, judicial oversight has some of the markings of a police patrol. It is regular and ongoing, and it involves large numbers of unremarkable agency decisions.

The ordinariness of judicial review relates to a second reason why it does not really serve as a form of fire alarm oversight in the context of high volume agency adjudication. To the extent that fire alarm oversight depends upon attracting the attention of Congress or the public at large, the regularity of court involvement interferes with the objective. We are unaware of any congressional hearings held during the past decade that court decisions in social security cases prompted, even as federal judges have fulminated about poor quality SSA decision-making.<sup>205</sup> If fire alarms ring all the time, then they seem less like alarms and more like background noise.

Finally, especially for the sorts of problems that courts are uniquely well-positioned to identify and to try to correct, effective judicial oversight of high volume agency adjudication is often not reactive and incident-driven, but requires judicial proactivity and extended engagement over time. Sometimes an appeal from a random ALJ or IJ order sounds the alarm over a large-scale matter whose significance a court immediately appreciates. When the BIA determined that someone seeking asylum based on her experience with female genital mutilation did not establish a risk of future persecution because the mutilation happened in the past,<sup>206</sup> the Second Circuit swiftly rebuked the agency for a “significant error[] in the application of its own regulatory framework.”<sup>207</sup> But an array of smaller bore but

---

204. *Chavarria-Reyes v. Lynch*, 845 F.3d 275, 280 (7th Cir. 2016) (Posner, J., dissenting).

205. We searched for congressional publications in the Proquest Congressional database using the search terms “federal /s (court or judge),” “social security,” “disability /s benefits,” and “Posner,” limiting our search to 2007 to 2017. The search yielded nothing suggesting a hearing or other oversight activity prompted by federal court opinions. “Posner” refers to Richard Posner. Because of Judge Posner’s stature and because of his high-profile criticism of disability benefits adjudication, his name should appear in oversight materials prompted by judicial criticism, had there been any.

206. *In re A-T-*, 24 I & N Dec. 296, 303–04 (B.I.A. 2007).

207. *Bah v. Mukasey*, 529 F.3d 99, 101 (2d Cir. 2008).

nonetheless important pathologies, such as problematic behavior by a single adjudicator or flaws in an agency's internal manual, can plague agency decision-making. Judicial awareness of these problems might sharpen only over time, and only as courts engage repeatedly with them.

C. *Problem-Oriented Oversight Through Judicial Review*

Judicial review of high volume agency adjudication does not fit the police patrol model either. The process relies upon third parties to identify and complain about flawed agency decision-making, which is a defining feature of fire alarm oversight. Courts do not proactively seek out adjudicator orders to review, as an auditor randomly sampling decisions to get an overall sense of the agency's performance might.<sup>208</sup> But an adjusted version of the police patrol metaphor works pretty well to describe the oversight role that courts can assume. "Problem-oriented policing"

posits that police should focus more attention on *problems*, as opposed to *incidents* . . . . Problems are defined either as collections of incidents related in some way (if they occur at the same location, for example) or as underlying conditions that give rise to incidents, crimes, disorder, and other substantive community issues . . . .<sup>209</sup>

Whereas "incident-driven," reactive policing focuses on the resolution of discrete incidents,<sup>210</sup> problem-oriented policing treats each incident as a datum for the identification of underlying factors that create crime and for the best possible responses.<sup>211</sup> Identifying underlying causes, not clearing arrests, is the goal.<sup>212</sup>

Table 2 describes definitional characteristics of fire alarm, police patrol, and problem-oriented oversight.

---

208. See, e.g., *Martinez v. Astrue*, 630 F.3d 693, 695 (7th Cir. 2011) (explaining that the government's inability to appeal as well as caution by claimants' lawyers in appealing make it impossible for courts to assess the error rate of administrative adjudications).

209. Gary Cordner & Elizabeth Perkins Biebel, *Problem-Oriented Policing in Practice*, 4 CRIMINOLOGY & PUB. POL'Y 155, 156 (2005).

210. ANTHONY A. BRAGA, PROBLEM-ORIENTED POLICING AND CRIME PREVENTION 9 (2d ed. 2010).

211. *Id.* at 10, 15.

212. Cordner & Biebel, *supra* note 209, at 156, 158.

Table 2. Models of Oversight Compared

Definitional Characteristic	Fire Alarm Oversight	Police Patrol Oversight	Problem-Oriented Oversight
<i>Initiator</i>	Third Party	Oversight Institution	Third Party
<i>Regularity of Oversight</i>	Episodic	Regular/Ongoing	Regular/Ongoing
<i>Goal of Oversight</i>	Problem Identification and Response	Overall Assessment of Performance	Problem Identification and Response
<i>Mode of Oversight</i>	Discrete Response to Incident	Audit of Numerous Agency Decisions	Discrete Response to Incident or to Pattern Gleaned from Review of Numerous Agency Decisions

When courts engage in problem-oriented oversight, they treat appeals as indicators of potential problems. Of course, many appeals simply result from adjudicator “error,” a word we use as a term of art. But “problems,” defined as systematic underlying pathologies in internal agency administration that afflict adjudication, can lurk among these flaws. The claimant or immigrant bringing the problem to a court’s attention may not know whether his case presents an error or a problem. Precisely the ordinariness of judicial review, or the continuing, routine engagement of courts with the agency’s decision-making, enables courts to distinguish problems from errors and respond appropriately.

*1. Errors.*—Agency adjudicators can produce flawed decisions for several reasons. Sometimes they simply err. The agency has adopted an acceptable interpretation of governing law. An acceptably competent adjudicator understands and applies this interpretation. But in the odd case, the adjudicator, as a mere mortal, happens to make an error. Perhaps amidst the six hundred pages of medical records in the claimant’s file, an ALJ overlooks the physician’s note that confirms a claimant’s alleged

symptoms.<sup>213</sup> Perhaps the IJ wrongly but not unreasonably treats a particular conviction as a “crime involving moral turpitude,” which requires the immigrant’s deportation.<sup>214</sup>

When an agency adjudicator errs, a reviewing court can correct the error but accomplish little more. By our definition of error, no underlying problem exists to address. Presumably, the ALJ would have decided the case better had she caught the physician’s note, and the case proceeded to federal court only because her mistake slipped past personnel at the Appeals Council. As we have already argued, this error correction offers a marginal justification for judicial review of high volume agency adjudication. To return to the metaphor, the error-correction function is like arresting a random lawbreaker, not ferreting out what underlying factors foster criminal activity.

2. *Problems.*—Flawed decisions result from problems, not mere error, in one of two situations. First, the agency may have adopted a bad policy. Second, the agency cannot or will not fix an entrenched decision-making pathology.

a. *Bad Policy.*—Agencies can adopt bad policies. The BIA’s erstwhile stance on female genital mutilation is an example. An instruction in a guidance document or manual that conflicts with governing precedent is another, albeit one more likely to fly under the radar and less likely to trigger a loud fire alarm.<sup>215</sup> However fine the mesh in its net, an internal appeals tribunal would never catch flawed adjudicator decisions when the shortcomings result from a bad policy because the tribunal has to abide by the policy as well. Thus, it would uphold an adjudicator’s decision following the policy as correct.

b. *Entrenched Pathology.*—A second type of problem results when the agency is unwilling or unable to correct an entrenched pathology that afflicts adjudicator decision-making. The threat of deliberate indifference to certain strains of adjudicator dysfunction lurks in the institutional DNA of agencies tasked with large numbers of claims or decisions to make. The number of cases decided is an easily administrable performance metric, but one that can reward decision-making that fares poorly by the harder-to-use measure of decisional quality.<sup>216</sup> If an agency sets production targets or quotas, as the EOIR and SSA do, it may find the temptation to ignore warning signs of

---

213. See *Oversight of Rising Social Security Disability Claims and the Role of Administrative Law Judges: Hearing Before the Subcomm. on Energy Policy, Health Care and Entitlements of the H. Comm. on Oversight & Gov’t Reform*, 113th Cong. 2 (2013) (statement of Sen. Tom Coburn) (“The average case has over 600 pages in it.”).

214. E.g., *Omargharib v. Holder*, 775 F.3d 192, 200 (4th Cir. 2014).

215. For examples of flawed guidance documents, see *Harris v. Astrue*, No. 3:09CV00260, 2010 WL 3909495, at \*5 (S.D. Ohio May 21, 2010); *Palaschak v. Astrue*, No. 08–CV–1172, 2009 WL 6315324, at \*11 (N.D.N.Y. Nov. 13, 2009).

216. See *supra* note 149 and accompanying text.



serious adjudicator dysfunction overwhelming.<sup>217</sup> Judge Daugherty, the Huntington ALJ, had a shockingly high allowance rate and decided astonishing numbers of cases. Together with the \$600 million in lifetime benefits he awarded,<sup>218</sup> these dubious achievements should have raised red flags in SSA headquarters.<sup>219</sup> Instead, notwithstanding a well-documented morale and management problem in the Huntington hearing office, the SSA, under pressure to keep a growing backlog at bay,<sup>220</sup> transferred 1,186 aged cases there between 2006 and 2011.<sup>221</sup> During this time, the SSA based its evaluations of ALJ performance solely on number of cases decided, with no adjustment for decisional quality.<sup>222</sup>

The Huntington episode did not trigger judicial review because the SSA generally cannot appeal when an ALJ grants benefits. But an agency focused

217. See Lisa D. Ordóñez et al., *Goals Gone Wild: The Systematic Side Effects of Over-Prescribing Goal Setting* 7–8 (Harvard Bus. Sch., Working Paper No. 09-083, 2009), <http://www.hbs.edu/faculty/Publication%20Files/09-083.pdf> [<https://perma.cc/9ACZ-3YAA>] (“Goals that are easier to achieve and measure (such as quantity) may be given more attention than other goals (such as quality) in a multi-goal situation.”).

218. Devlin Barrett & Damian Paletta, *Three Indicted for Alleged Social Security Fraud Scheme in Kentucky*, WALL STREET J. (Apr. 5, 2016), <https://www.wsj.com/articles/three-indicted-for-alleged-social-security-fraud-scheme-in-kentucky-1459867962> [<https://perma.cc/YGG6-AB4Z>].

219. STAFF OF H. COMM. ON OVERSIGHT & GOV’T REFORM, *supra* note 149, at 6.

220. For an indication of congressional interest in the disability claims backlog during the mid-2000s, see *Improve the Responsiveness and Oversight of the Hearings Process*, SOC. SEC. ADMIN., OFF. INSPECTOR GEN., <https://web.archive.org/web/20160809001205/https://oig.ssa.gov/audits-and-investigations/top-ssa-management-issues/social-security-disability-hearings-backlog?page=6> [<https://perma.cc/G6LJ-5E4K>] (flagging the hearing backlog as a “Top SSA Management Issue” and listing dozens of reports on the topic, including nearly 40 from 2005 to 2010).

221. STAFF OF S. COMM. ON HOMELAND SEC. & GOVERNMENTAL AFFAIRS, *supra* note 191, at 20–22.

222. STAFF OF H. COMM. ON OVERSIGHT & GOV’T REFORM, 113TH CONG., SYSTEMIC WASTE AND ABUSE AT THE SOCIAL SECURITY ADMINISTRATION 32–34 (2014). Another episode involved an ALJ in Harrisburg, Pennsylvania, who granted benefits to 2,285 people in 2007 alone. Although others in the agency criticized this ALJ, the agency’s chief ALJ praised him for “putting in incredible hours” and insisted that the ALJ “feels very committed to public service.” Brent Walth & Bryan Denson, *Paying Out Billions, One Judge Attracts Criticism*, OREGONIAN (Dec. 30, 2008), [http://www.oregonlive.com/special/index.ssf/2008/12/paying\\_out\\_billions\\_one\\_judge.html](http://www.oregonlive.com/special/index.ssf/2008/12/paying_out_billions_one_judge.html) [<https://perma.cc/2U9R-KWLW>]. The ALJ continued in his role as Hearing Office Chief Administrative Law Judge for a year and a half after his decision pattern made news. See *Social Security Administration Oversight: Examining the Integrity of the Disability Determination Appeals Process: Hearing Before the H. Comm. on Oversight and Gov’t Reform*, 113th Cong. 30 (2014) (statement of Charles Bridges, Administrative Law Judge, Social Security Administration) (stating that he continued in that role until June 2010, eighteen months after the *Oregonian* piece). He was removed from his leadership role in June 2010 because of how he administered the hearing office, but not explicitly because of his decision pattern. Appellees’ Supplemental Appendix at SA00166, *Bridges v. Comm’r of Soc. Sec.*, 607 F. App’x 168 (3d Cir. 2015) (No. 14-1580). Only in 2014 did the agency seem to take action to address the ALJ’s decision patterns. See *Bridges v. Comm’r of Soc. Sec.*, 607 F. App’x 168, 170 (3d Cir. 2015) (describing a determination made after FY 2013 that the ALJ’s decisions “did not comply with SSA standards” and subsequent action taken by his supervisors).

on numbers might just as well turn a blind eye to poor-quality decision-making that harms claimants or immigrants if the adjudicator decides a lot of cases.<sup>223</sup> The Atlanta immigration court decides cases in immigrants' favor at astonishingly low rates.<sup>224</sup> Persistent criticism for perceived bias against immigrants hounds Atlanta IJs,<sup>225</sup> and at least one Atlanta IJ has attracted a disproportionate number of formal complaints.<sup>226</sup> But, as an observer speculates, the EOIR has not taken significant steps at reform, perhaps because the Atlanta immigration court decides large numbers of cases.<sup>227</sup>

Entrenched pathologies might persist for reasons other than deliberate indifference, but ones equally baked into the institutional structure of agency adjudication. Agency adjudicators often enjoy employment protections that amount to a minor league version of life tenure.<sup>228</sup> The SSA cannot take disciplinary action against an ALJ based solely on how the ALJ decides

---

223. A Seattle-based IJ attracted scathing criticism from the Ninth Circuit and unfavorable public comment in 2002. *Paramasamy v. Ashcroft*, 295 F.3d 1047, 1048–49 (9th Cir. 2002); Chris McGann & Lise Olsen, *Controversial Immigration Judge Won't Be Transferred*, SEATTLE POST-INTELLIGENCER (Oct. 11, 2002), <http://www.seattlepi.com/news/article/Controversial-immigration-judge-won-t-be-1098273.php> [https://perma.cc/257W-LVHR]. She then moved to Los Angeles, where her decisions continued to garner unflattering attention. *See, e.g.*, *Smolnikova v. Gonzales*, 422 F.3d 1037, 1047 n.2 (9th Cir. 2005); *Rivera v. Ashcroft*, 387 F.3d 835, 841–42 (9th Cir. 2004); Pamela A. MacLean, *Immigration Judges Come Under Fire; Critics Say System Oversight Is Weak*, NAT'L L.J., Jan. 30, 2006. But she continued to decide cases, and she continued to garner severe criticism. *E.g.*, *Cruz Rendon v. Holder*, 603 F.3d 1104, 1111 n.3 (9th Cir. 2010). Asked why complaints against this IJ would prove futile, a prominent immigration attorney insisted that other IJs told him that the agency thought well of her work because “she clears a lot of cases.” John Roemer, *Jurist's Asylum Seeker Rulings Earn Rebuke*, L.A. DAILY J., Jan. 31, 2006 (quoting Robert Jobe).

224. *E.g.*, Ramji-Nogales et al., *supra* note 8, at 331 (indicating an asylum grant rate of 12% in Atlanta, compared to 40% overall).

225. *E.g.*, Letter from Hallie Ludsin, Emory Law Sch., et al., to Juan P. Osuna, Dir., Exec. Office for Immigration Review 5–6 (Mar. 2, 2017) (“[O]bservers [of Atlanta IJs] noted specific examples of concern where IJs made statements that indicated potential prejudice against immigrant respondents, or lacked the necessary patience, dignity, and courtesy required of IJs in immigration proceedings.”).

226. Between October 1, 2009, and March 31, 2010, at least five complaints were filed against IJ Pelletier. *See* Bryan Johnson, *Secret Identities of Immigration Judges Revealed*, AMJO LAW (Jan. 16, 2017), <https://amjolaw.com/2017/01/16/secret-identities-of-immigration-judges-revealed/> [https://perma.cc/HD4R-Z47C] (including a “modified key” that lists complaints against IJs by date filed). During this time period, eighty-seven complaints were filed against IJs nationwide. Executive Office for Immigration Review, Complaints Received Between Oct. 1, 2009, and Mar. 31, 2010 (on file with the authors). The EOIR employed 232 IJs in FY 2009. U.S. DEP'T OF JUSTICE, EXEC. OFFICE FOR IMMIGRATION REVIEW, FACT SHEET: EXECUTIVE OFFICE FOR IMMIGRATION REVIEW IMMIGRATION JUDGE HIRING INITIATIVE 3 (2010). The average IJ, in other words, received .38 complaints during the time period IJ Pelletier received five.

227. Jacqueline Stevens, *Lawless Courts*, NATION (Oct. 20, 2010), <https://www.thenation.com/article/lawless-courts/> [https://perma.cc/PLC9-AY45].

228. *See Continuing Oversight of the Social Security Administration's Mismanagement of Federal Disability Programs: Hearing Before the Subcomm. on Energy Policy, Health Care and Entitlements of the H. Comm. on Oversight and Gov't Reform*, 113th Cong. 27 (2013) (statement of Glenn E. Sklar, Deputy Commissioner, Social Security Administration) (explaining the SSA generally cannot issue strong discipline to ALJs such as furlough, suspension, or removal).

cases,<sup>229</sup> and its power to force ALJs to manage their cases in particular ways is tightly constrained.<sup>230</sup> An ALJ bears almost no risk of termination.<sup>231</sup> Indeed, the SSA believes that it cannot suspend an ALJ without pay, much less terminate him, until that ALJ has exhausted his appeals before the Merit Systems Protection Board (MSPB).<sup>232</sup> This extended process can create considerable delay.<sup>233</sup> After pleading guilty to a felony charge, for example, an ALJ who had sexually assaulted an employee in a hearing room during work hours while intoxicated received his salary for three more years until the MSPB had finally finished its review.<sup>234</sup>

Such protections, a (lesser) version of which IJs also enjoy,<sup>235</sup> give agency adjudicators a plausible claim to independence.<sup>236</sup> But they can lead to inertia or conflict avoidance within the agency and slow down or arrest efforts to respond to decision-making pathologies. Notwithstanding repeated federal judicial criticism of his performance,<sup>237</sup> for instance, one ALJ remained a hearing office chief administrative law judge until a class of 4,000 denied claimants filed a lawsuit against the SSA, alleging that due process violations systemically plagued his and several colleagues' case

229. See *Nash v. Bowen*, 869 F.2d 675, 681 (2d Cir. 1989) (stating that to “coerce ALJs into lowering reversal rates—that is, into deciding more cases against claimants—would, if shown, constitute in the district court’s words ‘a clear infringement of decisional independence’”).

230. *Soc. Sec. Admin. v. Butler*, No. CB-7251-14-0014-T-1, slip op. at 24–25 (M.S.P.B. Sept. 16, 2015); Emilia Sicilia, *Combating Biased Adjudication in Claims for Social Security Disability Benefits*, CLEARINGHOUSE COMMUNITY (May 2014), <http://povertylaw.org/clearinghouse/stories/sicilia> [<http://perma.cc/3AT4-GDSF>].

231. *Role of Social Security Administrative Law Judges: Joint Hearing Before the Subcomm. on Courts, Commercial and Admin. Law of the H. Comm. on the Judiciary and the Subcomm. on Soc. Sec. of the H. Comm. on Ways and Means*, 112th Cong. 46 n.10 (2011) (statement of Michael J. Astrue, Comm’r, Social Security Administration) (indicating that between 2008 and July 2011, the SSA tried to fire eight ALJs out of more than 1,000).

232. *Id.* at 46.

233. SOC. SEC. ADMIN., OFFICE OF THE INSPECTOR GEN., A-06-16-50026, AUDIT REPORT: ADMINISTRATIVE LEAVE USED FOR EXTENDED ABSENCES 3 (2017).

234. *Id.* at 3–4; Jillian Kay Melchior, *Social Security Disability Judge Got \$600,000 in Pay and Three Raises After Drunkenly Groping Colleagues*, NAT’L REV. (Jan. 26, 2017), <http://www.nationalreview.com/article/444272/sridhar-boini-social-security-judge-sexually-assaulted-colleagues-given-raises> [<https://perma.cc/Y5DF-4AMF>].

235. Stephen H. Legomsky, *Deportation and the War on Independence*, 91 CORNELL L. REV. 369, 373–74 (2006).

236. See Kent Barnett, *Against Administrative Judges*, 49 U.C. DAVIS L. REV. 1643, 1647 (2016) (describing ALJs as more independent than other administrative adjudicators who lack the same statutory protections).

237. *E.g.*, *Lazo-Espinoza v. Astrue*, No. 10–CV–2089, 2012 WL 1031417, at \*14 (E.D.N.Y. Mar. 27, 2012); *Bailey v. Astrue*, 815 F. Supp. 2d 590, 599–601 (E.D.N.Y. 2011); *Legare v. Comm’r of Soc. Sec.*, No. 08–CV–2180, 2010 WL 5390958, at \*3 (E.D.N.Y. Dec. 22, 2010); *Calderon v. Astrue*, 683 F. Supp. 2d 273, 278 (E.D.N.Y. 2010); *Gross v. Astrue*, No. 08–CV–578, 2010 WL 301945, at \*3 (E.D.N.Y. Jan. 15, 2010); *Ginsberg v. Astrue*, No. 05–CV–3696, 2008 WL 3876067, at \*16 (E.D.N.Y. Aug. 18, 2008).

management.<sup>238</sup> Only upon the lawsuit's filing did the SSA relieve the ALJ of his management role.<sup>239</sup>

3. *Distinguishing Errors from Problems.*—To succeed as overseers, courts have to be able to distinguish problems from errors. Sometimes the former are obvious. A sharp uptick in court remands suggests something more systematic afoot than idiosyncratic adjudicator error. When the SSA terminated disability benefits for hundreds of thousands of claimants in the early 1980s,<sup>240</sup> appeals flooded the courts, and the court remand rate jumped from 19% in 1980 to nearly 60% in 1984.<sup>241</sup> The SSA's problematic policies with regard to mental impairments and continuing-disability review quickly became obvious.<sup>242</sup> Likewise, if sufficiently awry, even a single flawed decision can suggest an entrenched pathology. The Ninth Circuit described an IJ's decision denying asylum in a 2005 case as "a literally incomprehensible opinion," "indecipherable," and "extreme in its lack of a coherent explanation,"<sup>243</sup> flaws that loudly signaled a troubled adjudicator.<sup>244</sup>

In many instances, however, problems manifest themselves less clearly. These are ones where the bad policy or the entrenched pathology is subtler, and thus demonstrates its faults only over time. The SSA provides ALJs with a digital template that generates boilerplate for decisions. Before 2012, this text included a poorly written paragraph that presented an ALJ's findings in a manner that suggested that the ALJ had improperly assessed the claimant's credibility.<sup>245</sup> This flawed boilerplate is an example of a bad policy. But it is one whose demerits as such—that is, as a policy and not a random error—would likely become evident only as courts saw the same boilerplate over and over again.

Courts catch problems of this scale by reviewing large numbers of cases, identifying patterns of flaws, and determining that something more than random error creates them. What follows is a highly stylized description of

---

238. Amended Class Action Complaint at 2, *Padro v. Astrue*, No. 11–CV–1788, 2013 WL 5719076 (E.D.N.Y. Oct. 18, 2013), [https://mhp.urbanjustice.org/sites/default/files/2011-05-04\\_Amended\\_Complaint.pdf](https://mhp.urbanjustice.org/sites/default/files/2011-05-04_Amended_Complaint.pdf) [<https://perma.cc/LD54-6LW9>].

239. Mosi Secret, *Rejected Disability Claims in Queens May Be Reheard*, N.Y. TIMES (Jan. 11, 2013), <http://www.nytimes.com/2013/01/12/nyregion/rejected-disability-claims-in-queens-may-be-reheard.html> [<https://perma.cc/7EXV-5A23>].

240. DERTHICK, *supra* note 131, at 5.

241. *Id.* at 145.

242. Levy, *supra* note 146, at 487.

243. *Recinos De Leon v. Gonzales*, 400 F.3d 1185, 1187, 1193–94 (9th Cir. 2005).

244. When interviewed about the Ninth Circuit's decision, the IJ insisted that "the arguments" from asylum claimants "were all the same." Chorney, *supra* note 60.

245. *Mascio v. Colvin*, 780 F.3d 632, 639 (4th Cir. 2015); *Bjornson v. Astrue*, 671 F.3d 640, 644–45 (7th Cir. 2012); *Freismuth v. Astrue*, 920 F. Supp. 2d 943, 952 (E.D. Wis. 2013); *Harvey v. Astrue*, No. CIV–10–393–SPS, 2012 WL 984299, at \*4 (E.D. Okla. Mar. 22, 2012); *Krusemark v. Astrue*, 725 F. Supp. 2d 829, 837 (S.D. Iowa 2010).

this process, one that no court of which we are aware actually uses. It owes a debt to a method the SSA has pioneered, using Appeals Council data to find problems in ALJ decision-making.<sup>246</sup> We believe it illuminates the mental steps courts proceed through as they identify problems. We present the method here to argue how courts should oversee high volume agency adjudication, and then defend their capacity to use it in Part IV.

The first step involves devising the proper classifications of potential problems. As with problem-oriented policing, broad classifications are “too heterogeneous” to yield much information about agency adjudication,<sup>247</sup> a claim we elaborate upon at length in our report.<sup>248</sup> Problem-oriented policing uses “highly nuanced and precise problem definitions.”<sup>249</sup> To understand what factors generate burglaries in Tucson, Arizona, for example, the police should not just keep track of “burglaries.” Instead they should also gather data on “burglaries in college dormitories,” “burglaries in neighborhoods with alleyways,” and so forth.

Problem-oriented oversight through judicial review should do the same. In the social security context, for example, courts should identify potential problems not as “remands,” or even “remands to the Brooklyn Hearing Office.” Rather, courts should develop categories that can identify flawed policies at the level of detail at which the agency crafts it, and they should use categories that can identify entrenched pathologies at the level at which they fester. The problems might be “treating source – opinion rejected without adequate articulation,” or “inadequate rationale for credibility finding.”<sup>250</sup> The entrenched pathology category might track decisions at the individual ALJ level, and certainly at the hearing office level.

To identify patterns and thus potential problems, courts could then use problem definitions to map data gathered from decisions. For any particular judicial review context the map would differ and depend on courts’ sense of where problems likely will come from and how they might materialize. Table 3 tracks reasons for remands from judges in the hypothetical District of East Dakota over a three-year period. It offers a simple illustration of how a federal district might organize data capturing arguments made and reasons given in social security cases.

---

246. Ray & Lubbers, *supra* note 143, at 1601–02; Letter from Michael J. Astrue, Comm’r, Soc. Sec. Admin., to Xavier Becerra, Ranking Member, Subcomm. on Soc. Sec., Comm. on Ways and Means, U.S. House of Representatives 6 (Dec. 5, 2012), [https://waysandmeans.house.gov/UploadedFiles/QFR\\_responses\\_MichaelAstrue\\_SS\\_6\\_27\\_12\\_BECERRA.pdf](https://waysandmeans.house.gov/UploadedFiles/QFR_responses_MichaelAstrue_SS_6_27_12_BECERRA.pdf) [<https://perma.cc/YRY5-V537>].

247. Michael D. Reisig, *Community and Problem-Oriented Policing*, 39 CRIME & JUST. 1, 7 (2010).

248. GELBACH & MARCUS, *supra* note 16, at 52–56.

249. Reisig, *supra* note 247, at 7.

250. *Top 10 Remand Reasons Cited by the Court on Remands to SSA*, SOC. SEC. ADMIN., [https://www.ssa.gov/appeals/DataSets/AC08\\_Top\\_10\\_CR.html#fy2015](https://www.ssa.gov/appeals/DataSets/AC08_Top_10_CR.html#fy2015) [<https://perma.cc/3RTQ-NZZW>].

Table 3. D.E.D. Remands as a Percentage of Appeals,  
by Reason Given, FY2014–2016

		Treating source— inadequate articulation	Inadequate rationale for credibility finding	Inadequate rationale given for weight, given consultative examiner's opinion	Mental disorder not adequately considered
Hearing Office No. 1	ALJ 1	0.33	0.4	0.22	0.29
	ALJ 2	0.25	0.25	0.4	0.33
	ALJ 3	0.6	0.62	0.65	0.74
Hearing Office No. 2	ALJ 4	0.3	0.33	0.22	0.4
	ALJ 5	0.58	0.64	0.56	0.72
	ALJ 6	0.37	0.19	0.29	0.42

Table 3 breaks down reasons for remands into more precise categories that may correspond to detailed policy decisions the agency might make. The SSA, for instance, might urge its ALJs to assess credibility in a particular manner, or to use a particular approach to considering mental impairments. These policy determinations should show up in arguments claimants make for remands and reasons courts give for ruling in their favor. Table 3 also recognizes the possibility that a particular ALJ might be deciding cases in a pathological way, or that a particular hearing office suffers from pathological management.

The district would then organize data on its judges' decisions, to see if they suggest any particular problems. The number in each of Table 3's cells is a fraction, indicating how often a court concludes that a particular ALJ's decisions contain particular flaws. The numerator represents the number of cases in which the court agrees that the ALJ's decision contains the flaw, and the denominator is the number of cases in which the claimant argues that the ALJ's decision contains the flaw. Organized thusly, the data yield a heat map that highlights potential problems. Table 3, for instance, indicates that ALJs 3 and 5 produce unusually high numbers of remands, regardless of the alleged flaw, and have done so consistently. Their decisions' high rate of failure across the board may suggest adjudicator dysfunction, and its persistence

over multiple years may indicate an entrenched pathology that the agency cannot or will not correct.

Table 4 gives an example of a heat map that indicates an entrenched pathology at the hearing office level.

Table 4. Hearing Office Pathology

		Treating source— inadequate articulation	Inadequate rationale for credibility finding	Inadequate rationale given for weight given consultative examiner's opinion	Mental disorder not adequately considered
Hearing Office No. 1	ALJ 1	0.6	0.6	0.4	0.6
	ALJ 2	0.52	0.63	0.55	0.33
	ALJ 3	0.58	0.4	0.48	0.59
Hearing Office No. 2	ALJ 4	0.3	0.33	0.22	0.33
	ALJ 5	0.4	0.27	0.38	0.4
	ALJ 6	0.37	0.19	0.29	0.42

The consistency with which the District of East Dakota finds fault with ALJ decisions from Hearing Office 1 suggests that the problem lies not with a single idiosyncratic ALJ but with some office-wide phenomenon. But the office-wide phenomenon is likely office-specific, because the ALJs from Hearing Office 2 enjoy markedly better success across the board. A bad policy should produce a heat map along the lines of what Table 5 illustrates.

Table 5. Bad Policy

		Treating source— inadequate articulation	Inadequate rationale for credibility finding	Inadequate rationale given for weight given consultative examiner's opinion	Mental disorder not adequately considered
Hearing Office No. 1	ALJ 1	0.33	0.25	0.28	0.65
	ALJ 2	0.25	0.33	0.34	0.66
	ALJ 3	0.4	0.44	0.32	0.59
Hearing Office No. 2	ALJ 4	0.36	0.37	0.22	0.64
	ALJ 5	0.18	0.19	0.25	0.7
	ALJ 6	0.41	0.3	0.11	0.62

Again, as far as we know, no court actually uses this method or something like it to identify problems with agency adjudication. But some courts have engaged in an impressionistic version of the method for social security and immigration cases. In a 2005 opinion, for example, the Third Circuit marshaled a number of examples from cases to document “a disturbing pattern of IJ misconduct” involving “intemperate or humiliating remarks” directed at immigrants.<sup>251</sup> The Second Circuit listed six previous instances when it had commented on a particular IJ’s inappropriate behavior in an opinion reversing the IJ for another episode of similar misconduct.<sup>252</sup> The Tenth Circuit identified repeated instances when it faulted the SSA for ALJ decisions that rely exclusively on boilerplate language for credibility discussions.<sup>253</sup> A district judge in Wisconsin came closer to what we recommend here when he buttressed a scathing critique of “a wholly dysfunctional administrative process within the Social Security Administration” with pages of statistics demonstrating the agency’s poor record before his court.<sup>254</sup>

251. *Wang v. Attorney Gen.*, 423 F.3d 260, 267–68 (3d Cir. 2005).

252. *Islam v. Gonzales*, 469 F.3d 53, 56–57 (2d Cir. 2006); *see also* *Huang v. Gonzales*, 453 F.3d 142, 150 (2d Cir. 2006) (also citing several previous cases rebuking IJ Chase’s conduct).

253. *Hardman v. Barnhart*, 362 F.3d 676, 679 (10th Cir. 2005).

254. *Freismuth v. Astrue*, 920 F. Supp. 2d 943, 945, 955–67 (E.D. Wis. 2013).



4. *Responding to Problems.*—Problem-oriented policing counsels for a variety of responses beyond the mere arrest of perpetrators to address patterns of criminal activity. A police department, for example, might deploy social workers alongside police officers when criminal activity involves mentally ill people. Hospitalization and treatment might be the interventions instead of arrest.<sup>255</sup> Congress as an oversight institution likewise can choose from an extensive menu of tools when it addresses problems within an agency.<sup>256</sup> The federal courts in contrast appear to lack remedial options beyond issuing remands. They seem confined to error correction, a form of reactive, incident-driven policing.

But courts in fact have several oversight tools at their disposal.<sup>257</sup> First, they can criticize agency adjudicators in terms calculated to cause consternation or shame. In a 2005 opinion, for instance, the Third Circuit denounced “[t]he tone, the tenor, the disparagement, and the sarcasm of the IJ” as “more appropriate to a court television show than a federal court proceeding.”<sup>258</sup> A district judge singled out an ALJ and insisted that his decision “shows a blatant disregard, not only of the legal standards, but of his obligations as a judicial officer and the basic rights and humanity of a vulnerable segment of our society, the disabled.”<sup>259</sup> Naming an IJ, the Second Circuit included an extended and detailed summary of the many errors he committed, including extensive quotations from the hearing he conducted, in

---

255. E.g., Cindy Chang, *Across L.A. County, Law Enforcement Looks for Resources to Deal with the Mentally Ill*, L.A. TIMES (June 20, 2016), <http://beta.latimes.com/local/california/la-me-ls-sheriff-mentally-ill-20160620-snap-story.html> [<https://perma.cc/2XDN-WEQC>]; see also Amy C. Watson et al., *Improving Police Response to Persons with Mental Illness: A Multi-Level Conceptualization of CIT*, 31 INT’L J.L. & PSYCHIATRY 359, 361 (2008) (describing the Crisis Intervention Team model created by the Memphis Police Department to respond to calls involving a person suffering from a mental illness).

256. WALTER J. OLESZEK, CONG. RESEARCH SERV., 7-5700, CONGRESSIONAL OVERSIGHT: AN OVERVIEW 9–14 (2010).

257. Christopher Walker has rigorously documented ways that federal courts reviewing agency adjudication do more to extend their influence than simply remand cases for further adjudication. Christopher J. Walker, *Referral, Remand, and Dialogue in Administrative Law*, 101 IOWA L. REV. ONLINE 84, 88 (2016) [hereinafter Walker, *Referral*]; Christopher J. Walker, *The Ordinary Remand Rule and the Judicial Toolbox for Agency Dialogue*, 82 GEO. WASH. L. REV. 1553, 1590–99 (2014). Courts use various tools, Professor Walker maintains, in part to “communicate to the agency specific—and oftentimes even systemic—problems identified by the reviewing court.” Walker, *Referral*, *supra*, at 90. We agree. The tools we describe here add to and build upon those Professor Walker has identified.

258. *Wang v. Attorney Gen.*, 423 F.3d 260, 269 (3d Cir. 2005); see also *Cham v. Attorney Gen.*, 445 F.3d 683, 686 (3d Cir. 2006):

The case now before us exemplifies the “severe wound . . . inflicted” when not a modicum of courtesy, of respect, or of any pretense of fairness is extended to a petitioner and the case he so valiantly attempted to present. Yet once again, under the “bullying” nature of the immigration judge’s questioning, a petitioner was ground to bits.

259. *Lazo-Espinoza v. Astrue*, No. 10–CV–2089, 2012 WL 1031417, at \*8 n.4 (E.D.N.Y. Mar. 27, 2012).

a 2006 opinion.<sup>260</sup> The Ninth Circuit reproduced an “incoherent” order by an IJ in full as an appendix to a scathing opinion, letting the IJ’s incompetence speak for itself.<sup>261</sup>

Courts can also exploit bureaucratic fault lines to force an agency to respond. Agencies that lack independent litigating authority, such as the SSA and the EOIR, control neither when they appeal to the federal circuits nor their advocacy before the federal circuits.<sup>262</sup> The DOJ takes a very conservative approach to what matters it wants to appear before the courts of appeals, wary of administrative law precedent that might affect the federal government’s litigating position trans-substantively.<sup>263</sup> Rather than risk an adverse appellate decision, the DOJ might pressure the EOIR or SSA to correct a problem instead. Another fault line involves the personnel who defend ALJ and IJ decisions in federal court. They are not the same as those who supervise agency adjudicators.<sup>264</sup> A DOJ lawyer may tire of defending questionable decisions that prompt hostile court reactions and request that the EOIR take some corrective action.<sup>265</sup> A court might threaten the agency’s lawyer with sanctions if the agency continues to insist on defending flawed decisions, or if the agency does not take steps to correct the problem.<sup>266</sup>

Courts can also adopt doctrines that raise the costs for agencies if they do not correct a problem. The Ninth Circuit applies something called the “credit-as-true” rule in social security cases. Until recently,<sup>267</sup> the most commonly identified flaw with ALJ decisions involved their failure to explain adequately why the opinion of the claimant’s treating physician did not establish the claimant’s disability.<sup>268</sup> In most circuits, courts will remand

---

260. *Huang v. Gonzales*, 453 F.3d 142, 149–50 (2d Cir. 2006).

261. *Recinos De Leon v. Gonzales*, 400 F.3d 1185, 1193–96 (9th Cir. 2005).

262. DAVID E. LEWIS & JENNIFER L. SELIN, ADMIN. CONFERENCE OF THE U.S.: SOURCEBOOK OF UNITED STATES EXECUTIVE AGENCIES 115–16, 116 n.296 (2012); GELBACH & MARCUS, *supra* note 16, at 144.

263. GELBACH & MARCUS, *supra* note 16, at 145–46; BERNARD ROSEN, HOLDING GOVERNMENT BUREAUCRACIES ACCOUNTABLE 127–28 (3d ed. 1998); Neal Devins & Michael Herz, *The Uneasy Case for Department of Justice Control of Federal Litigation*, 5 U. PA. J. CONST. L. 558, 572–73 (2003).

264. The DOJ’s Office of Immigration Litigation handles immigration appeals, and the SSA’s Office of General Counsel, along with the U.S. Attorney, litigates social security cases in the district courts.

265. *See, e.g., HUME, supra* note 126, at 24 (relaying interview comments that describe informally modifying procedures in response to a court decision).

266. A district judge in Wisconsin did just this in 2013. David Traver, *Warning of Sanctions for U.S. Attorney, up to and Including Disbarment*, TRAVER & TRAVER S.C., <http://www.ssacconnect.com/260-sanctions> [https://perma.cc/MWE3-SJGY].

267. The SSA replaced the treating-physician rule by regulation in January 2017. Revisions to Rules Regarding the Evaluation of Medical Evidence, 82 Fed. Reg. 5844, 5853 (Jan. 18, 2017) (to be codified at 20 C.F.R. pts. 404, 416).

268. *E.g., Garrison v. Colvin*, 759 F.3d 995, 1012–13 (9th Cir. 2014) (“[A]n ALJ errs when he rejects a medical opinion or assigns it little weight while doing nothing more than ignoring it,

cases with such treating-physician flaws. The ALJ gets another chance to explain why the treating physician's opinion does not merit deference.<sup>269</sup> In the Ninth Circuit, however, courts must "credit as true" treating physician evidence that the ALJ does not adequately discount.<sup>270</sup> If that evidence, taken as true, establishes the claimant's disability, the court will remand for the payment of benefits only and refuse to give the ALJ another crack at the case.<sup>271</sup> Particularly irritating to the SSA,<sup>272</sup> the credit-as-true rule raises the cost of ALJs' failure to grapple adequately with treating-physician evidence.

An additional tool dovetails with fire alarm oversight. Courts can draw media attention to what are otherwise obscure and ignored parts of the federal courts' docket with scathing commentary or by otherwise publicizing what can easily pass under the media's radar. Judicial commentary on adjudicator performance can buttress other advocates' calls or efforts for reform.<sup>273</sup> The complaint in *Padro v. Astrue*,<sup>274</sup> a class action filed in New York against the SSA, quoted from dozens of judicial opinions remanding claims to support allegations that some Queens Hearing Office ALJs systemically deprived claimants of due process.<sup>275</sup>

Finally, Article III judges can use their considerable prestige to pursue reform while off the bench. Disheartened by the problems that have plagued immigration adjudication,<sup>276</sup> Robert Katzmman of the Second Circuit first spearheaded a prominent study of immigrants' access to counsel,<sup>277</sup> then created a public interest law organization that represents thousands of immigrants in cases before IJs.<sup>278</sup> Margaret McKeown of the Ninth Circuit

---

asserting without explanation that another medical opinion is more persuasive, or criticizing it with boilerplate language that fails to offer a substantive basis for his conclusion.").

269. See ADMIN. CONFERENCE OF THE U.S., *supra* note 123, at 20 (describing Ninth Circuit's doctrine that often denies the ALJ a second chance as an exception to the rule).

270. *Id.*

271. *Garrison*, 759 F.3d at 1020.

272. See Revisions to Rules Regarding the Evaluation of Medical Evidence, 82 Fed. Reg. at 5859–60 ("In our view, the credit as true rule supplants the legitimate decisionmaking authority of our adjudicators, who make determinations or decisions based on authority delegated by the Commissioner. The credit as true rule is neither required by the Act nor by principles of due process.").

273. E.g., AM. BAR ASS'N COMM'N ON IMMIGRATION, REFORMING THE IMMIGRATION SYSTEM: PROPOSALS TO PROMOTE INDEPENDENCE, FAIRNESS, EFFICIENCY, AND PROFESSIONALISM IN THE ADJUDICATION OF REMOVAL CASES 2–19 (2010).

274. No. 11–CV–1788, 2013 WL 5719076 (E.D.N.Y. Oct. 18, 2013).

275. Amended Complaint, *supra* note 238, at 24–66.

276. Katzmman, *supra* note 96, at 6–7 (chronicling and lamenting serious issues in immigration adjudication).

277. STEERING COMM. OF THE N.Y. IMMIGRANT REPRESENTATION STUDY REPORT, ACCESSING JUSTICE: THE AVAILABILITY AND ADEQUACY OF COUNSEL IN IMMIGRATION PROCEEDINGS 1–2 (2011).

278. The organization is the Immigrant Justice Corps. For information about its case load, see *Our Story: Our Impact*, IMMIGRANT JUST. CORPS, <http://justicecorps.org/our-story/#impact> [https://perma.cc/88JC-4UHV].

helped kickstart a similar effort in San Diego,<sup>279</sup> as has Michael Chagares of the Third Circuit in New Jersey.<sup>280</sup>

#### IV. Evaluating Problem-Oriented Oversight

Problem-oriented oversight adds to the list of functions judicial review can play in the context of high volume agency adjudication. In Part II, we described institutional determinants that limited the contribution that any of the other functions, on its own, could make to the case for judicial review of high volume agency adjudication. Problem-oriented oversight strengthens judicial review's normative foundation only if it fares better by an analogous institutional measure.

Problem-oriented oversight depends upon private litigants being able to bring problems to the federal courts, the federal courts' capacity to identify and respond to problems, and the efficacy of those responses in terms of their ameliorative effect on agency policy and behavior. In several regards, these criteria resemble those that inform the choice between private enforcement through civil litigation, on one hand, and public administration through agency action on the other, as means for the implementation of a regulatory regime.<sup>281</sup>

The literature on private enforcement addresses problems that differ from the supervision of agency adjudication. An illustrative example is whether lawmakers should pursue automobile safety through agency enforcement, such as recalls, or through private civil litigation, such as tort lawsuits. But this scholarship helpfully identifies a number of institutional advantages and disadvantages that privately initiated litigation in generalist courts has, at least as it compares with some form of direct agency action. These considerations, or closely analogous ones, provide a useful blueprint to assess courts' capacity to engage in problem-oriented oversight. They suggest that the federal courts can perform this function successfully. Judicial review relies upon private litigants, those most directly affected, to bring flaws with agency decision-making to courts' attention. The process thus produces information about pathologies or bad policy efficiently. The federal courts' independence from the agencies under review and Congress can insulate their oversight from agency slack or political pressure. Finally, Article III courts have sufficient influence with agencies to push for ameliorative changes, and oversight focused on rooting out the sorts of problems we describe does not overtax their expertise.

---

279. Johanna S. Shiovoni, *ABA Immigration Justice Project Celebrates Its First Anniversary in San Diego*, NEWSL. (Fed. Bar Ass'n, San Diego Chapter, San Diego, Cal.), Spring 2009, at 6.

280. LORI A. NESSEL & FARRIN ANELLO, *DEPORTATION WITHOUT REPRESENTATION: THE ACCESS-TO-JUSTICE CRISIS FACING NEW JERSEY'S IMMIGRANT FAMILIES* i-ii (2016).

281. For a comprehensive list of the considerations implicated by this choice, see Stephen B. Burbank et al., *Private Enforcement*, 17 LEWIS & CLARK L. REV. 637, 662-71 (2013).

### A. *Efficiency*

The private enforcement of a regulatory regime through civil litigation enjoys several efficiency advantages over public administration. Private enforcement spares the expenditure of public resources on enforcement while leveraging the capacity of the private bar toward this end. It also relies upon those directly affected by the regulatory regime to trigger the enforcement process and thus likely produces information about the regime's implementation or lack thereof particularly readily.<sup>282</sup> The efficiency case for problem-oriented oversight through judicial review is less straightforward, but it probably favors it over other forms of agency oversight that do not rely upon private initiative.

*1. Resources.*—Private enforcement enjoys at least two types of resource advantages over public administration. First, the public bears only those direct costs that relate to the judiciary's involvement. Otherwise, the costs of enforcement are internalized by the plaintiff, the party seeking to benefit, and the defendant, the party that has allegedly violated the regime. Second, by delegating the law enforcement task to private lawyers, private enforcement multiplies the number of personnel involved in a regime's implementation without increasing the size of the federal bureaucracy.

Problem-oriented oversight through judicial review may not enjoy the first advantage as convincingly. Because the federal government is the defendant or appellee, it must foot its own defense costs and, at least for social security cases, pay EAJA fees when claimants obtain certain types of favorable outcomes.<sup>283</sup> The agency could invest these resources in, say, an expansive audit program if it did not have to litigate.

This sort of audit program, however, would require a politically dicey expansion of the federal bureaucracy. The SSA's program of pre-effectuation review offers a useful comparison. Each year, the agency's Division of Quality randomly selects a small percentage of ALJ decisions that are favorable to claimants, and thus cannot be appealed, for further review before notice of the favorable decision goes to the claimant. In FY 2015, for instance, the Division's 119 staff members reviewed about 4,500 decisions and identified concerns in approximately 900 of them.<sup>284</sup> The same year, the federal courts remanded 8,646 cases.<sup>285</sup> Keeping the rate at which Division staff members find flaws constant, assuming that each member's caseload remains fixed, and assuming that decisions denying and allowing claims

---

282. *Id.* at 662–64; Matthew C. Stephenson, *Public Regulation of Private Enforcement: The Case for Expanding the Role of Administrative Agencies*, 91 VA. L. REV. 93, 108 (2005).

283. On EAJA obligations, see, for example, *Shalala v. Schaefer*, 509 U.S. 292, 302 (1993).

284. SOC. SEC. ADMIN., OFFICE OF THE INSPECTOR GEN., A-12-15-50015, AUDIT REPORT: PRE-EFFECTUATION REVIEWS OF FAVORABLE HEARING DECISIONS 1–2, 4 n.14 (2017).

285. SOC. SEC. ADMIN., *supra* note 85.

contain errors with the same frequency, the Division would have to expand by more than 1,000 staff members to catch the same number of mistakes as the federal courts do. By delegating much of the problem-identification task to private litigants and federal judges, judicial review spares the SSA this immense bureaucratic expansion.<sup>286</sup>

Problem-oriented oversight through judicial review is not necessarily as resource-friendly as private enforcement, although the politics of bureaucratic expansion may make its costs easier for Congress to swallow. But the case for judicial review requires more. In Part II, we questioned the value of court-based error correction on opportunity-cost grounds. The same concern warrants discussion here: if the resources invested in judicial review were spent instead on agency adjudication, would fewer problems arise in the first place?

On this score, the distinction between problems and errors makes the case for problem-oriented oversight stronger than that for error correction. Errors may result because an overworked ALJ does not have time to review a lengthy set of medical records thoroughly, or because an overextended IJ cannot probe an immigrant's story deeply enough. Logically, if the ALJ or IJ had more time, as a lower case load might permit, she would make fewer such errors. If the agency adopts a bad policy, however, an increase in adjudicator resources will do nothing to decrease the number of problematic adjudicator decisions. All decisions that comport with the policy, whether issued by a harried adjudicator or a relaxed one, will suffer.

The same outcome likely obtains when problems result from entrenched agency pathologies. If an SSA hearing office is mismanaged or suffers from bad morale, the addition of a new ALJ or two, or the hiring of three new decision writers, likely will not have a dramatic ameliorative effect. If an IJ harbors bias against immigrants, or if an ALJ thinks that most claimants are lazy ne'er-do-wells, a 10% caseload reduction is unlikely to change her mind. Excessive caseloads may deepen a pathology's entrenchment,<sup>287</sup> but a positive correlation does not necessarily or even often exist between caseloads and pathologies. The SSA's Miami Hearing Office, for instance,

---

286. Presumably, the SSA's Office of General Counsel (OGC) could shrink significantly if the agency did not have to defend its decisions in the federal courts. Presently, OGC has about 600 lawyers. *Regional Chief Counsel (Atlanta)*, USAJOBS (Sept. 5, 2017), <https://www.usajobs.gov/GetJob/PrintPreview/478378400> [<https://perma.cc/RS5F-ZCRF>]. If one assumes that each OGC attorney spends five-sixths of his or her time on federal court appeals, an end to judicial review could enable the SSA to downsize OGC by 500 lawyers. An investment of these resources in Division of Quality staff would still require a net increase of 500 personnel.

287. See Marcia Coyle, *Burnout, Stress Plague Immigration Judges*, NAT'L L.J. (July 13, 2009), <http://www.law.com/nationallawjournal/almID/1202432173266/?Sreturn=20170929151055> [<https://perma.cc/C8ME-QRT9>] ("[H]igh levels of burnout and stress may make it difficult for immigration judges to recognize trauma in the refugees who come before them.").

suffers from management and morale problems,<sup>288</sup> even though its productivity ranked it nearly last among the country's 163 offices in FY 2017.<sup>289</sup> In FY 2012, the year claimants filed *Padro v. Astrue*, the Queens Hearing Office decided fewer cases per day per ALJ than those from any other hearing office in the country.<sup>290</sup>

*2. Information Production.*—Another efficiency concern relates closely to the resources consideration. Private enforcement compares favorably to public administration because it relies on those with the best information, the injured parties, to identify misconduct and initiate a response. A version of this advantage is one of the chief arguments in favor of fire alarm oversight.<sup>291</sup> Rather than proactively audit an agency itself, Congress can more efficiently monitor agency performance if third parties bring misconduct to its attention.

Judicial review unquestionably brings problems with agency decision-making to the fore more cheaply than some sort of internal agency auditing process can. Depending upon how court access gets structured, barriers to judicial review can select for cases that are most likely to involve flawed decisions.<sup>292</sup> As discussed in Part II, hurdles for social security claimants can discourage a lot of potential appeals, and presumably those with strong claims are more likely to tough it out. Lawyers who represent social security claimants, to mention one barrier, get paid either by contingency or through EAJA fees, both of which require a claimant victory in federal court. Such hurdles should ensure that, of the appeals that get filed in federal court, many involve flawed ALJ decisions. Some of these decisions will involve errors and not problems, to be sure, and thus problem-oriented oversight succeeds only if courts can reliably distinguish between the two categories. But the subset is unlikely to involve a large number of correct decisions the way a

---

288. SOC. SEC. ADMIN., OFFICE OF THE INSPECTOR GEN., A-12-15-50041, WORKLOAD OVERSIGHT IN THE MIAMI HEARING OFFICE (2016).

289. *National Ranking Report by ALJ Dispositions per Day per ALJ FY 2017*, SOC. SEC. ADMIN., [https://www.ssa.gov/appeals/DataSets/archive/04\\_FY2017/04\\_September\\_Disposition\\_Per\\_Day\\_Per\\_ALJ\\_Ranking\\_Report.html](https://www.ssa.gov/appeals/DataSets/archive/04_FY2017/04_September_Disposition_Per_Day_Per_ALJ_Ranking_Report.html) [<https://perma.cc/5UAX-NGT2>].

290. *National Ranking Report by ALJ Dispositions per Day per ALJ FY 2012*, SOC. SEC. ADMIN., [https://www.ssa.gov/appeals/DataSets/archive/04\\_FY2012/04\\_September\\_Disposition\\_Per\\_Day\\_Per\\_ALJ\\_Ranking\\_Report.pdf](https://www.ssa.gov/appeals/DataSets/archive/04_FY2012/04_September_Disposition_Per_Day_Per_ALJ_Ranking_Report.pdf) [<https://perma.cc/3XKK-HQJ5>].

291. McCubbins & Schwartz, *supra* note 183, at 168.

292. The incentives that fuel appeals or barriers that limit them may be more complicated in other contexts. Some have argued, albeit with little empirical basis, that overly lax policies of granting stays of removal pending review have incentivized immigrants to file meritless appeals. Kagan et al., *supra* note 108, at 688, 692–94, 722–23. If so, then the pool of appeals before the circuits will include plenty of reasonable IJ decisions. In contrast, robust evidence suggests that detention discourages appeals. MILLER ET AL., *supra* note 103, at 131–32. Given that the immigrant can leave detention if she abandons her appeal and accepts removal, the fact that she remains incarcerated increases the likelihood that the IJ's decision includes an error or resulted from a problem.

random audit might, and thus the system operates efficiently to bring problems to courts' attention.

The SSA's Division of Quality example is again illustrative. From 2011 to 2015, the Division of Quality randomly selected 1.4% of ALJ decisions allowing benefits for pre-effectuation review.<sup>293</sup> In 80% of instances, the division "effectuated" the case with no further action taken, suggesting that it found grounds for concern in only one out of five cases it reviewed.<sup>294</sup> Over the same period, the federal courts remanded 43% of social security appeals.<sup>295</sup> Although the comparison between the two rates is not straightforward, it suggests, however crudely, that properly incentivized private litigants identify flawed decisions, and thus generate information for oversight, more efficiently than a random audit can.

### B. Independence

The efficiency case for problem-oriented oversight through judicial review contrasts it with something like an audit, a method that relies on agency personnel proactively searching for flaws in adjudicator decision-making. But agencies can engage in their own version of problem-oriented oversight through an appeals system. Internal appellate review places the onus on the private litigant to come forward and thus should generate information about agency performance more efficiently than a randomized audit, if not as markedly so as Article III review.<sup>296</sup>

Problem-oriented oversight through internal appellate review only works if appellate personnel within the agency can catch problems and respond to them successfully. In recent years the Appeals Council and the BIA have attracted criticism for inconsistent and sometimes arbitrary decision-making.<sup>297</sup> This perceived problem surely results, at least in part, from institutional determinants, including a paucity of time and resources. If Appeals Council adjudicators have to decide up to twelve cases per day, then their capacity to detect and respond to problems likely suffers. But the institutional case for Article III review does not depend upon whether these

---

293. SOC. SEC. ADMIN., OFFICE OF THE INSPECTOR GEN., *supra* note 284, at 1.

294. *Id.* at 2.

295. For data on the percentage of remands from 2011–2015, see SOC. SEC. ADMIN., *supra* note 57.

296. One reason why internal agency appellate review might not generate information as efficiently is that the barriers to appeal are lower. To appeal an ALJ's decision, for example, a claimant typically files little more than a three-page, often boilerplate letter identifying grounds for reversal. GELBACH & MARCUS, *supra* note 16, at 28. Moreover, someone who appeals to federal court has already appealed and lost within the agency, and thus has been pursuing her appeal for longer and more doggedly than those who have only appealed within the agency.

297. E.g., David Hausman, *The Failure of Immigration Appeals*, 164 U. PA. L. REV. 1177, 1180–81 (2016); see also GELBACH & MARCUS, *supra* note 16, at 28 ("The last two digits of a claimant's social security number—not, say, the hearing office from which an appeal comes—determines the branch to which an appeal goes.").



critiques are accurate or not. Article III review promises several independence advantages that internal appellate review lacks.

The literature on private enforcement identifies independence as an important advantage privately initiated litigation enjoys over direct agency action. Public administration can suffer from “agency slack,” or “the tendency of government regulators to underenforce certain statutory requirements because of political pressure, lobbying by regulated entities, or the laziness or self-interest of the regulators themselves.”<sup>298</sup> A concern in times of divided government that the President might steer agencies away from Congress’s regulatory objectives prompted the sharp increase from the 1960s onwards in the number of statutes delegating enforcement to private litigation.<sup>299</sup> Several analogous influences can interfere with an agency’s self-oversight. Review in Article III courts insulates oversight from these pressures and enjoys an institutional advantage for this reason.

*1. Agency Interests.*—An agency may be tempted to oversee its adjudicators in a manner that casts their performances in the best possible light or that avoids internal conflict. In 2012, for example, the DOJ’s Inspector General faulted the EOIR for measuring its own performance in a manner that “overstate[d] the actual accomplishments of” immigration courts.<sup>300</sup> The EOIR used a method for counting case completions that exaggerated IJ productivity, and it assessed efforts to meet timeliness goals in a way that did not capture how long immigrants actually had to wait to get their cases decided.<sup>301</sup> A quality-review system at the Board of Veterans’ Appeals samples one out of every twenty decisions by veterans’ law judges (VLJs) to look for flaws.<sup>302</sup> A decision is considered flawed only if no reasonable VLJ would have issued the decision under scrutiny, not if the reviewer thinks the case was actually decided incorrectly.<sup>303</sup> This threshold may avoid conflict with VLJs, who might resent second-guessing by personnel of less bureaucratic stature. But it does not come close to predicting how well VLJ opinions will fare on appeal.<sup>304</sup>

The self-interest problem can taint oversight through internal appellate review as well. The SSA uses the Appeals Council “agree” rate as an

---

298. Stephenson, *supra* note 282, at 110.

299. SEAN FARHANG, *THE LITIGATION STATE: PUBLIC REGULATION AND PRIVATE LAWSUITS IN THE U.S.* 216–17 (2010).

300. U.S. DEP’T OF JUSTICE, OFFICE OF THE INSPECTOR GEN., *supra* note 65, at i.

301. *Id.* at i–ii.

302. U.S. GOV’T ACCOUNTABILITY OFFICE, GAO-05-655T, *VA DISABILITY BENEFITS: BOARD OF VETERANS’ APPEALS HAS MADE IMPROVEMENTS IN QUALITY ASSURANCE, BUT CHALLENGES REMAIN FOR VA IN ASSURING CONSISTENCY* 7 (2005).

303. *Id.* at 9–10.

304. *See id.* at 6 (reporting that the Court of Appeals for Veterans Claims reversed or remanded 88% of the VLJ decisions it reviewed).

indicator of ALJ performance, mining internal appeals for information in a manner similar to what we describe in Part III. A rising agree rate indicates improved ALJ performance, or so the logic goes.<sup>305</sup> But if the Appeals Council's review becomes more deferential, then a rising agree rate indicates nothing at all about improved ALJ performance. Under these conditions, not only does internal appellate review function less successfully as an oversight mechanism, it can also affect other agency oversight methods that rely upon information generated by the appellate tribunal.<sup>306</sup>

Finally, an agency may simply not want to oversee itself, even if it can glean information efficiently through internal appellate review. This tendency is all but guaranteed when it comes to problems of flawed policy. If the SSA instructs ALJs to use certain flawed text for discussions of credibility, the Appeals Council will not fault ALJs for doing so, and the problem will not show up in Appeals Council decision patterns. The SSA has mined Appeals Council data to identify and root out some entrenched decision-making pathologies, the second type of problem. But, as far as we know, the EOIR has not used BIA decisions for this purpose.<sup>307</sup> In fact, as far as we know, neither the EOIR nor the DOJ's Inspector General has assessed the quality of IJ decision-making using BIA data. Certainly neither has embarked upon an effort to identify and respond to problems commensurate with the campaign against pathologies in immigration cases the federal courts of appeals have waged.

2. *Political Independence.*—Related to agency self-interest is politics' looming influence. An agency might not prioritize problem-oriented oversight, even if internal appeals offer it an opportunity to do so efficiently, if such oversight is politically inexpedient. An agency might align its self-

---

305. See SOC. SEC. ADMIN., OFFICE OF THE INSPECTOR GEN., A-12-16-50106, AUDIT REPORT: OVERSIGHT OF ADMINISTRATIVE LAW JUDGE DECISIONAL QUALITY 1–2 (2017) (stating that “managers use agree rate results as well as other quality reviews to ensure ALJ decisionmaking is consistent and accurate”); Ray & Lubbers, *supra* note 143, at 1604–06 (associating a declining “rate at which the Council remands to the hearing level” with “quality improvement”).

306. The SSA's Inspector General, like all inspectors general, enjoys protections that enable it to examine the SSA's decision-making without interference from the rest of the agency. See Rachel E. Barkow, *Overseeing Agency Enforcement*, 84 GEO. WASH. L. REV. 1129, 1176 & n.225 (2016) (noting some of the “various institutional design protections” that assist investigations by inspectors general); Shirin Sinnar, *Protecting Rights from Within? Inspectors General and National Security Oversight*, 65 STAN. L. REV. 1027, 1035–36 (2013) (discussing the “broad investigative powers” of inspectors general). But the SSA Inspector General does not generate raw data for assessment purposes on its own; it instead relies upon what the agency itself assembles. In a recent report on decisional quality, for example, it relied exclusively on the Appeals Council's agree rate as an ALJ performance measure. SOC. SEC. ADMIN., OFFICE OF THE INSPECTOR GEN., *supra* note 284, at 1.

307. AM. BAR ASS'N COMM'N ON IMMIGRATION, *supra* note 273, at 2-21–2-22. We submitted a Freedom of Information Act request to the EOIR asking for information about its quality assurance programs. We did not receive any information in response that indicated that the EOIR has used BIA information for this purpose.

policing with what it perceives as Congress's preferences. Congress can insist upon this alignment by enacting legislation requiring the agency to focus on particular problems.<sup>308</sup>

The agency may prioritize certain forms of oversight over others, even in the absence of legislation directing it to do so, to minimize conflict with Congress. Starting in 2011, roughly at the same time as the Huntington scandal, the SSA began to use Appeals Council data to identify problematic ALJs for "focused reviews."<sup>309</sup> Of the first fifty ALJs selected, thirty were identified because they had allowance rates that exceeded 75%.<sup>310</sup> By FY 2013, the number of high-allowance-rate ALJs had dropped precipitously,<sup>311</sup> a fact the agency's Chief Administrative Law Judge emphasized when she insisted at a Senate Committee hearing that "quality is improving."<sup>312</sup> But the number of low-allowance-rate ALJs, whose decisions are especially likely to generate court remands, ticked up slightly during the same period.<sup>313</sup> All of this happened as the SSA endured intense Congressional scrutiny for its perceived profligacy with benefits.<sup>314</sup>

In recent years, Congressional oversight of immigration policy has emphasized enforcement.<sup>315</sup> President Trump's first budget blueprint proposed that Congress authorize the EOIR to hire seventy-five new IJs, insisting that doing so would help to "combat[] illegal entry and unlawful presence in the United States."<sup>316</sup> In light of such pressures, the likelihood that the EOIR will prioritize oversight that looks for problems disadvantaging immigrants seems low.<sup>317</sup>

Congress's formal power to oversee the federal courts notwithstanding,<sup>318</sup> its attempts to exercise this power have been modest compared with its scrutiny of federal agencies.<sup>319</sup> Moreover, the federal

308. Section 845(a) of the Bipartisan Budget Act of 2015, for example, requires the SSA to report on its efforts to combat fraud and prevent improper payment. Bipartisan Budget Act of 2015, Pub. L. No. 114-74, § 845(a), 129 Stat. 584, 618.

309. STAFF OF H. COMM. ON OVERSIGHT & GOV'T REFORM, *supra* note 149, at 6.

310. *Id.* at 13 & n.31.

311. *Hearing on SSDI Abuse*, *supra* note 196, at 131 fig.1.

312. *Id.* at 130.

313. *Id.* at 131 fig.1.

314. See STAFF OF H. COMM. ON OVERSIGHT & GOV'T REFORM, *supra* note 149, at 40–41 (reporting with detail on individual ALJ adjudicators and high allowance rates).

315. ANTJE ELLERMANN, STATES AGAINST MIGRANTS: DEPORTATION IN GERMANY AND THE UNITED STATES 106 (2009).

316. OFFICE OF MGMT. & BUDGET, EXEC. OFFICE OF THE PRESIDENT, AMERICA FIRST: A BUDGET BLUEPRINT TO MAKE AMERICA GREAT AGAIN 30 (2017).

317. On the susceptibility of immigration adjudication to political pressure, see AM. BAR ASS'N COMM'N ON IMMIGRATION, *supra* note 273, at 2–24.

318. Todd David Peterson, *Congressional Investigations of Federal Judges*, 90 IOWA L. REV. 1, 33–39 (2004).

319. For instance, while Congress has created inspectors general for a number of agencies, including the DOJ and the SSA, efforts to do the same for the federal courts have failed repeatedly.

courts' diverse docket insulates them from some sort of politicized retaliation should their decisions in agency appeals tend to skew in one manner or another. Congress could always respond to a pattern of decisions it dislikes by altering the federal courts' jurisdiction or changing a standard of review. But short of such focused legislation, Congress is unlikely to use another sanction, like a budget cut, to pressure the federal courts because doing so will adversely affect other, more privileged, areas of their docket.

### C. Capacity

Our critique of judicial review's regulative function questions the capacity of courts to force agencies to abide by precedent. Judicial efforts to engage in problem-oriented oversight warrant the same scrutiny, although what information presently exists indicates that courts may succeed in prodding agencies to respond to their diagnoses of certain problems. The literature on private enforcement suggests two other reasons to question judges' capacity to administer regulatory regimes: their inexperience and the limited geographic reach of their decisions. Neither is a concern for court-based problem-oriented oversight.

*1. Efficacy of Judicial Interventions.*—The most obvious objection to judicial review's oversight function involves its efficacy. Neither the EOIR nor the SSA mines court remands for information that might help its adjudicators improve. One might expect agencies to act with similar indifference when courts respond more aggressively to perceived problems.

A federal court can all but ensure that an agency will respond if it uses extreme measures, such as Rule 11 sanctions, injunctive relief, or an approval of a class action settlement requiring changes.<sup>320</sup> Courts rarely do so, however.<sup>321</sup> Still, the difference between an ordinary court remand and the sort of opinion a court might issue when addressing a problem gives reason to think that the latter can influence agency operations.<sup>322</sup>

The agency can fully comply with an ordinary court decision if an adjudicator conducts the proceedings on remand in accordance with the court's instructions. If the court demands nothing more, it cannot fault the

---

Casey C. Sullivan, *Is It Time for an Inspector General of the Federal Courts?*, FINDLAW (July 8, 2015), [http://blogs.findlaw.com/federal\\_circuit/2015/07/is-it-time-for-an-inspector-general-of-the-federal-courts.html](http://blogs.findlaw.com/federal_circuit/2015/07/is-it-time-for-an-inspector-general-of-the-federal-courts.html) [<https://perma.cc/8VFP-FFZQ>].

320. The SSA recently settled a class action, for example, that requires it to give claimants who were evaluated by a particular consulting physician a chance to seek benefits again. Plaintiffs' Motion & Memorandum of Points & Authorities in Support of Motion for Final Approval of Class Action Settlement Agreement at 4, *Hart v. Colvin*, 310 F.R.D. 427 (N.D. Cal. 2015) (No. 15-cv-00623-JST).

321. *E.g.*, HUME, *supra* note 126, at 39 (describing courts' sparing use of sanctions and preference for lighter reprimands).

322. For similar optimism, see Walker, *Referral*, *supra* note 257, at 89–90.

agency for treating the remand as a one-off and not a source of constructive criticism. A decision concluding that “the ALJ failed to give specific and legitimate reasons for discounting [a treating physician’s] opinion,”<sup>323</sup> for instance, obliges the SSA to do no more than ensure the ALJ does so on remand, regardless of whether the ALJ’s hearing office is dysfunctional or if the ALJ routinely struggles with such evidence. A vast linguistic gulf separates this remand from a decision like *Freismuth v. Astrue*,<sup>324</sup> where the district judge denounced disability adjudication as a “wholly dysfunctional administrative process” and threatened the SSA with “very deep trouble” if it didn’t take steps to fix observed problems.<sup>325</sup> In response to the decision, the U.S. Attorney for the Eastern District of Wisconsin insisted that his office had “been very much in conversation and communication – some of it quite productive – with” the SSA.<sup>326</sup>

Robert Hume concluded his empirical study of agency responsiveness to courts with the finding that “words” in opinions like *Freismuth* “seem to matter,”<sup>327</sup> for several reasons. First, “[w]hen opinion language leaves agencies little room to maneuver, administrators might change their policies to avoid sanctions and maintain favorable relationships with judges.”<sup>328</sup> As repeat players, agencies know that they risk angering a judge who will surely decide appeals going forward if they ignore clear instructions to change course. While an angry judge could take out her anger on only a small number of cases relative to the agency’s overall case load, agencies value their “reputational capital” and “credibility” with the federal courts and do not want to dissipate them.<sup>329</sup> Perhaps for this reason, the Department of Justice has long had a policy of initiating an investigation any time a federal court of appeals identifies an IJ by name in an opinion.<sup>330</sup>

Second, Hume suggests that a clear, strongly worded opinion can empower certain individuals within an agency who may prefer the sort of

---

323. *Penoza v. Berryhill*, No. C15-1825-RAJ, 2017 WL 1532667, at \*4 (W.D. Wash. Apr. 28, 2017).

324. 920 F. Supp. 2d 943 (E.D. Wis. 2013).

325. *Id.* at 945, 954.

326. Jane Pribek, *Federal Judges Fired Up Over Social Security Cases*, WIS. L.J. (Mar. 11, 2013), <http://wislawjournal.com/2013/03/11/federal-judges-fired-up-over-social-security-cases/> [https://perma.cc/Y39T-2PX4].

327. HUME, *supra* note 126, at 126.

328. *Id.* at 78–79.

329. *Id.* at 116.

330. U.S. DEP’T OF JUSTICE, OFFICE OF PROF’L RESPONSIBILITY, POLICIES AND PROCEDURES 2 (2015) (on file with authors); Immigration Prof., *L.A. Immigration Judge Under Fire*, LAW PROFESSOR BLOGS NETWORK (Dec. 20, 2009), <http://lawprofessors.typepad.com/immigration/2009/12/la-immigration-judge-under-fire.html> [https://perma.cc/LJ8W-Q98V]. The SSA is more vague about how it responds to judicial criticism, but insists that it “carefully analyzes Federal court decisions” and “value[s] the courts’ perspective . . .” Marilyn Odendahl, *Disability Denials Draw Criticism*, INDIANALAWYER.COM (Apr. 6, 2016), <https://www.theindianalawyer.com/articles/39934-disability-denials-draw-criticism> [https://perma.cc/7XK2-36JS].

policy adjustment the court counsels relative to those who favor the status quo.<sup>331</sup> Others have documented this “destabilization effect” within federal agencies that judicial opinions can produce.<sup>332</sup> Perhaps agency officials have ignored a general counsel’s recommendation that adjudicators use different language when discussing someone’s credibility. The right sort of judicial opinion faulting the agency for its credibility boilerplate can give the general counsel significant leverage to insist upon a policy change.<sup>333</sup>

Third, as Hume reports, “[r]esearch on administrative behavior . . . emphasizes that administrators are professionals who take their work seriously and try to do what is right.”<sup>334</sup> Agency officials may feel obliged out of a sense of professional obligation to respond when courts give unambiguous and strongly worded feedback.<sup>335</sup> This assumption, that agency personnel see themselves as professionals trying to discharge their mission as successfully as possible, underlies many of the SSA’s efforts to improve ALJ performance.<sup>336</sup> It might also explain why congressional oversight is often effective.<sup>337</sup> Congress rarely passes legislation when a fire alarm rings. An agency may worry about its budget appropriation, but investigatory committees rarely have budgetary powers, and appropriations are a clumsy, blunt tool to use to insist upon specific change.<sup>338</sup> Maybe congressional oversight works because agencies want to do the right thing. If so, court pressure can have the same effect.

2. *Expertise*.—A standard critique of private enforcement compares courts unfavorably to agencies as generalists lacking in sufficient expertise to administer a regulatory regime optimally.<sup>339</sup> One version of this critique challenges judicial review’s oversight function on grounds that courts cannot diagnose problems with adjudication as expertly as agencies can. The charge has force in two instances. First, a judicial attempt to force agencies into large-scale procedural changes of the sort that could dramatically upend settled agency practice should give pause. As Adrian Vermeule argues, “[t]he federal judicial system is not set up, not equipped, to engage in a sustained

---

331. HUME, *supra* note 126, at 75–76.

332. Hal G. Rainey, *What Motivates Bureaucrats?*, 12 J. PUB. ADMIN. RES. & THEORY 303, 305 (2002) (reviewing MARISSA MARTINO GOLDEN, *WHAT MOTIVATES BUREAUCRATS? POLITICS AND ADMINISTRATION DURING THE REAGAN YEARS* (2001)); Charles F. Sabel & William H. Simon, *Destabilization Rights: How Public Law Litigation Succeeds*, 117 HARV. L. REV. 1015, 1020 (2004).

333. HUME, *supra* note 126, at 76.

334. *Id.* at 8–9.

335. *Id.* at 113.

336. Ray & Lubbers, *supra* note 143, at 1598.

337. Beerman, *supra* note 178, at 121–22.

338. Kriner, *supra* note 180, at 784–85.

339. *See supra* note 120.

course of synoptic institutional engineering.”<sup>340</sup> But, as Vermeule also argues, the federal courts, aware of their institutional limitations, have largely surrendered control over fundamental matters of procedural design to agencies.<sup>341</sup>

The expertise critique also has some bite when courts fail to appreciate that agency adjudicators have to optimize how they conduct their proceedings under significant constraints. Although some federal judges have a decent sense of the limits under which agency adjudicators labor,<sup>342</sup> others may be surprisingly unaware of adjudicator caseloads and their inadequate support.<sup>343</sup> Attempts to micromanage how adjudicators manage cases deserve criticism, as federal judges may not understand how resource inadequacies constrain the process agency adjudicators can afford.<sup>344</sup>

Most of the problems we have discussed in this Article, however, require neither a deep appreciation for immutable determinants that require adjudicators to act in certain ways nor an omniscient eye for large-scale procedural design. The fact that IJs decide 1,000 cases per year does not excuse IJ bias against categories of immigrants. The SSA has to ask ALJs to decide 500–700 cases per year; flaws in the credibility boilerplate the agency has ALJs insert into their decisions does not help them work through their dockets more quickly. Properly conducted, problem-oriented oversight should operate as a form of arbitrary and capricious review, a type of oversight that permits the agency to continue in a particular procedural vein if it has a plausible reason to do so.

A second version of the expertise critique is unique to judicial review of high volume agency adjudication. This data gathering and analysis we describe in Part III may seem far afield from core judicial competencies and may beg the question of whether courts deciding one case at a time can assemble information usefully from individual appeals that can accurately indicate problems.

To a certain extent, our method merely illustrates the sort of thinking that a judge should engage in to identify patterns and spot problems. A court does not have to assemble precisely the heat map we describe. Indeed, court competencies probably enable a district or circuit clerk’s office to develop an even more sophisticated approach to problem identification. Some courts already assemble some of the sort of information that a problem-oriented

---

340. ADRIAN VERMEULE, *LAW’S ABNEGATION* 115 (2016).

341. *Id.* at 123–24; cf. Paul Verkuil, *Meeting the Mashaw Test for Consistency in Administrative Decision-Making*, in *ADMINISTRATIVE LAW FROM THE INSIDE OUT: ESSAYS ON THEMES IN THE WORK OF JERRY L. MASHAW* 239, 246 (Nicholas R. Parrillo ed., 2017).

342. Cf. *Chavarria-Reyes v. Lynch*, 845 F.3d 275, 280 (7th Cir. 2016) (Posner, J., dissenting) (“[The Immigration Court] may well owe its dismal status to severe underfunding by Congress . . .”).

343. GELBACH & MARCUS, *supra* note 16, at 5–6.

344. See *supra* notes 158–162 and accompanying text.

court would harvest from individual appeals. The Ninth Circuit does so for all cases, not just one category or another. There, a staff attorney reviews each appeal once it is fully briefed, judges its complexity, and prepares a “case inventory” that identifies the issues the appeal raises. The issues get entered into a searchable database to enable the Ninth Circuit to track it along with cases raising similar issues.<sup>345</sup>

The data analysis that problem-oriented judicial review requires should likewise pose little challenge. The patterns courts can identify in the data should prompt them to look at relevant appeals in a different light, but they should not react mechanistically to some statistical anomaly as conclusive proof of a problem. The SSA looks for outliers in ALJ decision data as guides to where it needs to investigate further.<sup>346</sup> A trend’s emergence in court data should likewise further investigation, albeit of the sort that a court can undertake. Perhaps the fact that courts remand an IJ’s claims involving immigrant credibility at an unusually high rate should signal to a judge that she take a hard look at a particular appeal for signs of IJ bias. Judges should not automatically remand a case involving mental impairments, much less pen some screed on bad SSA policy, simply because remand data indicate a sharp uptick across ALJs and hearing offices for cases involving mental impairments. But such indications would signal to judges to pay particular attention to how the agency describes applicable policy in such cases.

3. *Geographical Dispersion.*—The literature on private enforcement cites the judiciary’s geographic dispersion as a comparative disadvantage.<sup>347</sup> A federal agency can administer a regulatory regime uniformly, subjecting the regulated entity to a consistent set of constraints nationwide. In contrast, regulation through private tort litigation, for example, subjects the defendant to different obligations in different places.

Geographic dispersion creates a somewhat different difficulty for problem-oriented oversight through judicial review. When the problem is one of a flawed policy, a court decision faulting the agency for its adoption suffers the same limitations as one attempting to regulate the agency through precedent. The agency, motivated by a felt obligation to administer a single policy nationally and concerned about adjudicator inconsistency, might resist making any changes in response to judicial chastisement.

When, however, a problem involves an entrenched decision-making pathology, the federal judiciary’s geographic dispersion is often a feature, not a bug. Provided that venue rules require that decisions from a particular set

---

345. Harry Pregerson & Suzanne D. Painter-Thorne, *The Seven Virtues of Appellate Brief Writing: An Update from the Bench*, 38 SW. L. REV. 221, 223 (2008).

346. Ray & Lubbers, *supra* note 143, at 1594–95.

347. See, e.g., Burbank et al., *supra* note 281, at 667–68 (blaming a “decentralized” judiciary as part of why private enforcement regimes lead to “fragmented and incoherent policy”).



of adjudicators go consistently to a particular set of judges,<sup>348</sup> a geographically dispersed system of judicial review will better ensure that pathologies discoloring adjudication in a particular immigration court or a particular hearing office come to a federal judge's attention. Most appeals from disability-benefits decisions rendered by ALJs in the Tucson Hearing Office get filed in the District of Arizona.<sup>349</sup> A District of Arizona judge will see decisions by the same ALJ repeatedly and certainly will review decisions from the same hearing office. If, however, all social security appeals were to proceed in a single national social security court, the chances are slim that one of its judges would see multiple appeals from the same ALJ or that one of its judges would develop a feel for a problem arising at one of the SSA's 166 hearing offices. If cases are randomly assigned, then a lot of time could pass before one of the national court's judges saw the same ALJ's name on an appealed decision, or even the same originating hearing office. A judge on this national court would be more likely to mistake a problem for an error.

### Conclusion

The standard justifications for judicial review of high volume agency adjudication are unsatisfying. Institutional clashes interfere with the corrective, regulative, and critical functions the federal courts attempt to serve, rightly prompting doubt that the benefits courts create when they discharge these functions exceed judicial review's costs. Problem-oriented oversight, suffering from fewer of these institutionally determined limitations, creates additional benefits. When added to the mix, the contributions courts make when they ferret out problems tip the balance in favor of judicial review.

We recognize that the costs and benefits of judicial review are difficult to quantify with precision.<sup>350</sup> Reasonable people may disagree with the empirical assertions we make about how courts can act and how agencies might respond. Even so, an understanding of problem-oriented oversight is important for at least two reasons. First, as future scholars and policy makers rethink judicial review of high volume agency adjudication, they should

---

348. On venue choices for social security cases, see 42 U.S.C. § 405(g) (2012). On venue for immigration cases, see 8 U.S.C. § 1252(b)(2) (2012).

349. A Tucson ALJ will most likely decide cases involving Tucson claimants. See HALLEX, *supra* note 158, at I-2-0-70 ("The [hearing office] will generally process all requests for hearing (RH) for claimants residing in that area."). Appeals from Tucson claimants to the federal courts most likely will go to the District of Arizona. See 42 U.S.C. § 405(g) ("Such action shall be brought in the district court of the United States for the judicial district in which the plaintiff resides . . .").

350. We thank Andy Coan for helping us to formulate this concluding thought.

measure courts' capacities to identify and help fix problems as they assess the value of all the contributions courts can make. Second, and perhaps more important, judicial review is here to stay, at least for the time being. As long as it remains so, courts can maximize the value they add to agency adjudication by engaging in problem-oriented oversight.