

University of Pennsylvania Carey Law School

Penn Law: Legal Scholarship Repository

Faculty Scholarship at Penn Law

7-28-2016

A Study of Social Security Disability Litigation in the Federal Courts


Jonah B. Gelbach

University of Pennsylvania Carey Law School

David Marcus

University of Arizona

Follow this and additional works at: https://scholarship.law.upenn.edu/faculty_scholarship

 Part of the [Administrative Law Commons](#), [Civil Procedure Commons](#), [Courts Commons](#), [Disability Law Commons](#), [Law and Society Commons](#), [Policy Design, Analysis, and Evaluation Commons](#), [Public Administration Commons](#), [Public Economics Commons](#), [Public Policy Commons](#), [Social Welfare Commons](#), and the [Social Welfare Law Commons](#)

Repository Citation

Gelbach, Jonah B. and Marcus, David, "A Study of Social Security Disability Litigation in the Federal Courts" (2016). *Faculty Scholarship at Penn Law*. 1668.

https://scholarship.law.upenn.edu/faculty_scholarship/1668

This Report is brought to you for free and open access by Penn Law: Legal Scholarship Repository. It has been accepted for inclusion in Faculty Scholarship at Penn Law by an authorized administrator of Penn Law: Legal Scholarship Repository. For more information, please contact PennlawIR@law.upenn.edu.



Administrative Conference of the United States

**A STUDY OF SOCIAL SECURITY LITIGATION IN THE
FEDERAL COURTS**

Final Report: July 28, 2016

Jonah Gelbach
University of Pennsylvania

David Marcus
University of Arizona

This report was prepared for the consideration of the Administrative Conference of the United States. The opinions, views and recommendations expressed are those of the authors and do not necessarily reflect those of the members of the Conference or its committees, except where formal recommendations of the Conference are cited.

Acknowledgments

We owe thanks to a number of people. Gisselle Bourns, Stephanie Tatham, Matt Wiener, Reeve Bull, and Dan Sheffner at the Administrative Conference of the United States gave us invaluable guidance and feedback on all stages of this project. Gerald Ray, David Foster, Rainbow Lloyd, and Ben Gurga of the Social Security Administration answered innumerable questions and provided access to essential information and data. Nearly 150 people who regularly participate in various stages of social security disability claims adjudication agreed to be interviewed and gave generously of their time.

Table of Contents

Introduction	4
Part I. District Court Review and its Discontents	9
A. The Discomfort of District Court Review.....	9
B. Judicial Review’s Persistence	11
Part II. The Path of a Disability Claim	16
A. The Disability Claims Adjudication Process.....	16
1. The Initial Application	16
2. The Hearing Office	18
3. The Appeals Council.....	27
4. Federal Court	30
B. The Numbers.....	35
C. Quality Assurance	36
Part III. The Federal Court Remand Rate and What It Means	44
A. The Remand Rate in Perspective	45
B. The Strictness of District Court Review.....	50
1. The Quality of Decisions Appealed	51
2. The Substantial Evidence Standard of Review	56
3. The Missed Opportunity of Appellate Review	67
C. The Quality of Agency Adjudication	71
1. Different Goals and Resources	72
2. Different Legal Commitments	74
3. Different Perspectives.....	77
Part IV. Variations Across Districts	81
A. A Quantitative Analysis of District-Level Variation in Remands.....	86
1. The Nature of the Problem of District-Level Variations.....	87
2. Are Outlier Judges Important?.....	91
3. A Circuit-Level Component to District-Level Variation	95
B. A Quantitative Analysis Using Multiple Regression	98
1. Hypotheses and Variables.....	100
2. Statistical Model and Results.....	102
C. Using ODAR Regions to Capture Intra-Agency Factors	110
D. Qualitative Evidence of Varying Input Quality	117
E. Reflections on District-Level Variations.....	123
Part V. The Procedural Governance of Social Security Litigation	127
A. Variations in Social Security Procedure.....	128
B. Problems With Disuniformity	132
C. Problems With Particular Procedures.....	136
1. Joint Submissions.....	136
2. Simultaneous Briefing	139
3. Pre-Motion Letters	140
4. Oral Argument	141
5. Deadlines, Page Limits, and Extensions	141
Part VI. Recommendations and Suggestions	143
A. Recommendations for ACUS Consideration	143
1. Independent Litigating Authority.....	143
2. Social Security Rules Enabling Legislation.....	148
3. Social Security Rules	152
4. Education and Communication.....	159
5. Addressing Circuit Variation	163
B. Suggestions for the Agency.....	167

Introduction

A person who has sought and failed to obtain disability benefits from the Social Security Administration (“SSA” or “the agency”) can appeal the agency’s decision to a federal district court. In 2015, nearly 20,000 people did just this. Even though claims pass through multiple layers of internal agency review, many of them return from the federal courts for even more adjudication. A claimant’s experience in the federal courts, however, differs considerably from district to district around the country. If she files suit in Brooklyn, New York, a particular set of procedural rules will apply, and she will litigate before a district judge in a district where seventy percent of claimants’ appeals are remanded. If she files suit in Little Rock, Arkansas, she will encounter a different set of procedures, a magistrate judge will most likely decide her case, and her case will be heard in a district where claimants win much less often—only about twenty percent of the time.

The adjudication of disability claims within the agency has received relentless attention from Congress, government inspectors general, academic commentators, and others. Social security litigation in the federal courts has not weathered the same scrutiny. Federal judges adjudicate many fewer social security cases than the more than 500,000 decisions that administrative law judges (“ALJs”) render annually. Measured by caseloads alone, the federal judiciary’s importance to the implementation of American disability policy is modest. Nonetheless, an investigation into the nature of and challenges posed by this litigation is overdue. Through case law and remands, the federal judiciary exercises an outsized influence on disability claims adjudication. Social Security cases add a hefty amount to the federal courts’ workload, contributing seven percent of filings nationwide, and claimants file suit in increasing numbers. Finally, while inconsistency in ALJ decision-making has drawn attention, seemingly erratic decision making in district courts has only recently prompted investigation.

This Report, prepared for the Administrative Conference of the United States, is the first comprehensive study of social security disability litigation in the federal courts.¹ Our charge included the investigation of three questions:

- What factors explain why claimants prevail so often when they appeal to the federal courts, even after multiple layers of review within the agency?
- What factors explain variations in remand rates among various federal districts?
- How does the litigation of disability appeals vary from district to district?

We combined traditional legal research with quantitative analysis to find answers. The agency supplied us with a substantial amount of data containing information on outcomes of disability appeals in the federal courts. We also relied on docket report information drawn from PACER, which one of us has used in other research,² as well as publicly available data. To supplement and guide our quantitative inquiry, and following the lead of past studies of disability claims adjudication,³ we interviewed about 150 people involved in the disability claims adjudication process. Our interview subjects included federal judges and their law clerks, ALJs, agency officials and staff, claimant representatives,⁴ and Department of Justice personnel.

This investigation revealed one obvious fact: federal judges know little about the path social security claims follow from initial filing to their chambers. Indeed, a number of judges and clerks we interviewed reported mistaken impressions of what happens inside the agency, to

¹ A prior study documented and analyzed inconsistencies in district court outcomes for social security litigation. See Harold J. Krent & Scott Morris, *Inconsistency and Angst in District Court Resolution of Social Security Disability Appeals*, 67 HASTINGS L.J. 367 (2016). This study investigates causes of inconsistencies in district court remand rates, a subject we address in Part IV. Our Report addresses a number of other topics as well. Moreover, while some of our empirical findings comport with theirs, our quantitative investigation into district-level disparities adds new statistical inquiries. Our data set also differs. See *infra* note 470.

² For a general description of these data, see Jonah B. Gelbach, *Material Facts in the Debate Over Twombly and Iqbal*, 68 STAN. L. REV. ___ (forthcoming 2016) (Mar. 18, 2015, draft), at 17. For a discussion of the specific docket information we used in this Report, see Data Appendix.

³ JERRY L. MASHAW ET AL., SOCIAL SECURITY HEARINGS AND APPEALS 7 (1978).

⁴ “Claimant representative” is a term used to refer to those who represent social security claimants. A claimant representative does not have to be an attorney to represent someone in agency proceedings. See 20 C.F.R. §§ 404.1705 and 416.1505. All of the claimant representatives we interviewed, however, litigate in the federal courts and are attorneys.

such an extent that judicial misapprehensions may actually affect outcomes in some instances.⁵ After introducing the challenges of judicial review in Part I, we seek to fill judges' knowledge gap with an extensive summary of disability claims adjudication in Part II. Its contents should provide federal judges with a badly needed introduction to the mysterious process that generates so many of the cases they decide.

In Part III, we analyze the question of why claimants win so many of their federal court appeals. While ALJs could always improve their performance, and while the federal courts may have unrealistic expectations for decisional quality, we ultimately conclude that many claimant wins result from a largely unavoidable clash between these two institutions. The agency and the federal courts have conflicting goals, resources, priorities, and legal commitments. Even if both institutions are performing adequately, we argue, federal courts will continue to rule against the agency in a large number of cases.

We turn to the problem of inconsistencies in district court decision-making in Part IV. The institutional conflicts we describe in Part III are identical nationwide, and yet the rate at which claimants secure remands varies considerably from district to district. Our quantitative inquiry into this puzzle produced three primary findings:

- District and magistrate judges tend to march in lockstep within districts. Districts with one judge who remands a lot of cases to the agency tend to have other judges who do so as well. Very few individual judges have decision patterns that depart significantly from what their district colleagues produce.
- Circuit boundaries are associated with a good deal of district-level variation. For example, the fact that the Eastern District of New York remands more cases than the Southern District of Florida seems to be significantly related to the fact that, over all,

⁵ We interviewed 24 federal judges. In many instances their law clerks participated extensively in the discussions. We also interviewed 24 ALJs, 14 decision writers, 12 claimant representatives, 12 agency personnel who work for the SSA's Office of Disability Adjudication and Review, 31 lawyers who work for the SSA's Office of General Counsel, and 3 Department of Justice officials. In addition, we participated in a meeting of representatives from various U.S. Attorney's Offices.

districts in the Second Circuit remand a greater share of cases than do districts in the Eleventh Circuit.

- A number of factors – judicial ideology, the degree of a district’s urbanization, the assignment of cases to district versus magistrate judges, ALJ case loads, and others – have little association with case outcomes.

These results are consistent with a separate hypothesis. Decisions coming out of different hearing offices may vary by quality, in ways that could affect district court decision-making. We could not confirm that this was so with the data available to us. We nonetheless report the results of an intensive qualitative investigation of three federal districts and the hearing offices within them, to explain this hypothesis and to urge the agency to inquire into it further.

We address the procedural governance of social security litigation in Part V. Parties litigate social security appeals pursuant to a dizzying array of local rules, district-wide orders, and individual judge preferences. Whether the agency has to answer a complaint, what sort of merits briefs the parties have to file, and a host of other procedural questions have answers that differ considerably from district to district and sometimes from judge to judge. These procedural differences have few benefits, create inefficiencies, and impose other costs. We describe problems caused by many of the procedural variations.

Part VI contains suggested reforms. We take as an assumption that Congress will not significantly adjust the current system of social security adjudication or litigation. Also, our recommendations are informed by our sense of the limited mandate that ACUS has for suggesting changes to internal agency processes. These recommendations, then, are modest. They aim to improve the efficiency of social security litigation and enable litigants to explain their positions better to federal courts. We do not believe that these recommendations would necessarily cause a significant drop in the national remand rate if they were implemented. As we argue, this rate is determined by a complex assortment of factors, including the two institutions’

goals, resources, priorities, and commitments.⁶ But these recommendations could reduce needless expense and perhaps produce more uniformity in case law. Part VI closes with suggestions for the agency to consider as it continues its efforts to improve the quality of social security disability claims adjudication.

⁶ Further, there are good reasons to believe that claimant behavior would adjust to changes in the appeals process. This means the quality composition of cases appealed to the district courts can be expected to change when the process itself changes. It is possible that the overall remand rate would not change much—or perhaps even would rise—if such compositional changes were great enough.

Part I. District Court Review and its Discontents

A. The Discomfort of District Court Review

In 1939, Congress gave the federal district courts jurisdiction to review SSA decisions on claims for old-age and survivors benefits.⁷ The agency set up what remains for all intents and purposes the current structure for benefits adjudication the following year, one that culminates with federal court review.⁸ When Congress created the modern disability benefits system in 1956,⁹ claimants availed themselves of the pre-existing structure to appeal denied claims.¹⁰

At least formally, district court review of denied claims is unremarkable. Unless a statute specifies otherwise, appeals of agency actions, including decisions by administrative law judges, go to district courts by default.¹¹ Health care providers seeking Medicare reimbursement from the U.S. Department of Health and Human Services¹² and farmers seeking federal benefits from the U.S. Department of Agriculture can appeal to district courts when dissatisfied with agencies' decisions.¹³

In practical terms, however, district court review of disability benefits appeals is singular. First, no other type of appeal from an administrative agency generates anywhere near the volume of litigation for district courts that disability claims do. During the twelve-month period ending June 30, 2014, for example, plaintiffs filed 515 discrimination lawsuits against the United States as a defendant, some portion of which presumably came on appeal from the Merit Systems

⁷ 42 U.S.C. § 405(g); Social Security Act Amendments of 1939, Sen. Rpt. No. 734, 76th Cong., First Sess., July 7, 1939, at 52.

⁸ Irving Ladimer, *Hearing and Review of Claims and Wage-Record Cases Under Old-Age and Survivors Insurance*, 3 SOC. SEC. BULL. 21, 22-24 (1940).

⁹ E.g., MARTHA DERTHICK, POLICYMAKING FOR SOCIAL SECURITY 304-306 (1979); Charles I. Schottland, *Social Security Amendments of 1956: A Summary and Legislative History*, SOC. SEC. BULL., Vol. 19, No. 9, Sept. 1956, at 3.

¹⁰ E.g., George C. Longshore, *The Social Security Disability Insurance Program – An Example of the Necessity for Judicial Review of Administrative Decisions*, 25 ALA. LAW. 282, 284-286 (1964).

¹¹ *Watts v. Securities & Exchange Comm'n*, 482 F.3d 501, 505 (D.C. Cir. 2007); 3 RICHARD J. PIERCE, JR., ADMINISTRATIVE LAW TREATISE § 18.2, at 1680 (5th ed. 2010).

¹² 42 U.S.C. § 1395ff(b)(1)(A); 42 U.S.C. § 1395w-22(g)(5).

¹³ 7 U.S.C. § 6999.

Protection Board.¹⁴ Medicare reimbursement appeals do not merit a line item in the *Statistical Tables for the Federal Judiciary* and presumably account for some of the 987 “other statutory actions” involving the United States as a defendant.¹⁵ During this same time period, disability claimants filed 18,676 cases.¹⁶ Immigrants challenging adverse agency decisions do flood the federal courts with cases, but they file their roughly 6,000 appeals in the circuit courts.¹⁷ District courts would review agency actions infrequently but for their social security docket.

Second, the district courts rarely sit in an appellate capacity, a point helpfully stressed to us by one of the federal judges we interviewed.¹⁸ They decide a steady diet of habeas corpus petitions, but habeas review involves a highly specialized area of law that differs in crucial respects from other appellate or appellate-type litigation.¹⁹ District judges do entertain a substantial number of appeals from bankruptcy judges, but disability appeals dwarf these cases by a factor of about nine to one.²⁰

The position district courts occupy in the disability determination process is not only singular. Our impressionistic sense from the interviews we conducted with federal judges is that many do not relish their social security docket.²¹ Frustration with the agency sometimes surfaces

¹⁴ Statistical Tables for the Federal Judiciary – June 2014, Table C-2, at 2, at <http://www.uscourts.gov/statistics-reports/statistical-tables-federal-judiciary-june-2014> (last visited Nov. 11, 2015).

¹⁵ *Id.*

¹⁶ *Id.* at 4. This total only includes “Disability Insurance” and “Supplemental Security Income” cases.

¹⁷ 6,927 appeals from agency decisions were taken to the circuits in FY 2014, 86% of which were appeals from the Board of Immigration Appeals. See www.uscourts.gov/statistics-reports/us-courts-appeals-judicial-business-2014 (last visited Nov. 11, 2015).

¹⁸ Federal Judge 24 at 2. All interview subjects were promised anonymity and are referred to in this report by generic identifiers that describe them by their positions. Page numbers refer to the authors’ interview notes, all of which are on file with the authors.

¹⁹ RICHARD H. FALLON ET AL., HART AND WECHSLER’S THE FEDERAL COURTS AND THE FEDERAL SYSTEM 1345 (4th ed. 1996).

²⁰ Statistical Tables for the Federal Judiciary – June 2014, Table C-2, at 2, at <http://www.uscourts.gov/statistics-reports/statistical-tables-federal-judiciary-june-2014> (last visited Nov. 11, 2015).

²¹ Federal Judge 3 at 1 (“No one likes to do social security cases.”); Federal Judge 4 at 2 (describing social security cases as “the bane of our existence”); Federal Judge 5 at 3 (“Most Article III judges dread these cases.”); Federal Judge 6 at 4 (ranking disability claims appeals with prison litigation and habeas work as “the three most undesirable parts of my docket”); Federal Judge 7 at 3 (describing disability claims appeals as a “horribly ill fit for the skill set of Article III judges and clerks”); Federal Judge 2 at 4 (reporting that, while he likes disability claims appeals, his

in opinions.²² Conversely, federal judges have earned persistent criticism for their role in the disability claims process, going back to the very start of the current system for claims adjudication.²³ Calls for the replacement of district court review surface regularly.²⁴ For many of the agency personnel we interviewed, district courts remain the round hole for the square peg of disability claims adjudication.²⁵ Federal judges seem to have little understanding of what happens in hearing offices, some ALJs complained to us.²⁶ Others have criticized federal judges for their lack of sympathy for or understanding of the SSA's obligation to administer a national program consistently.²⁷

B. Judicial Review's Persistence

Is judicial review worth it? In 1978, Jerry Mashaw and a distinguished group of colleagues published a path-breaking study of the disability claims adjudication process that surveyed and critiqued all of its aspects.²⁸ Their study, while nearly forty years old, remains the

colleagues disagree with him by a factor of 2:1). A couple of federal judges we interviewed felt otherwise. *See* Federal Judge 1 at 4 (“Judges who single out social security irritate me. It’s their job.”); Federal Judge 2 at 4 (“This area of law is a passion.”).

²² *E.g.*, Hughes v. Astrue, 705 F.3d 276, 279 (7th Cir. 2013); Hardman v. Barnhart, 362 F.3d 676, 679 (10th Cir. 2004); Freismuth v. Astrue, 920 F. Supp. 2d 943, 954 (E.D. Wis. 2013); Batista v. Colvin, Civ. No. 13-4185, 2014 U.S. Dist. LEXIS 80576, at *6 (E.D.N.Y. June 11, 2014); Taylor v. Astrue, Civ. No. 07-3469, 2008 WL 2437770, at *5 (E.D.N.Y. June 17, 2008).

²³ Landon H. Rowland, *Judicial Review of Disability Determinations*, 52 GEO. L.J. 42, 86 (1962) (insisting that social security appeals “provided the federal courts with a severe test of their ability to deal effectively with an administrative process”); Robert G. Dixon, Jr., *The Welfare State and Mass Justice: A Warning from the Social Security Disability Program*, 1972 DUKE L.J. 681, 702 (criticizing district courts for their “marked tendency . . . to convert the program into a humanitarian unemployment program for any worker whose unemployment had some demonstrable basis in physical or mental trauma”); Richard E. Levy, *Social Security Disability Determinations: Recommendations for Reform*, 1990 BYU L. REV. 461, 511-12; R. Shep Melnick, *Administrative Law and Bureaucratic Reality*, 44 ADMIN. L. REV. 245, 256 (1992); Paul R. Verkuil & Jeffrey S. Lubbers, *Alternative Approaches to Judicial Review of Social Security Disability Cases*, 55 ADMIN. L. REV. 731, 739-741 (2003).

²⁴ *E.g.*, SOCIAL SECURITY ADVISORY BOARD, IMPROVING THE SOCIAL SECURITY ADMINISTRATION’S HEARING PROCESS, Sept. 2006, at 17; Verkuil & Lubbers, *supra* note 23 at 778; Levy, *supra* note 23 at 512;

²⁵ *E.g.*, ALJ 2 at 2 (describing the district courts as speaking a different language when they address social security issues).

²⁶ *E.g.*, ALJ 3 at 4 (complaining that federal judges do not understand the pressures ALJs are under); ALJ 4 at 5 (complaining that federal judges do not understand “our process”).

²⁷ *E.g.*, Harold H. Bruff, *Coordinating Judicial Review in Administrative Law*, 39 U.C.L.A. L. REV. 1193, 1210 (1992).

²⁸ JERRY L. MASHAW ET AL., SOCIAL SECURITY HEARINGS AND APPEALS 145-146 (1978). Mashaw’s co-authors were Charles J. Goetz (University of Virginia), Frank Goodman (University of Pennsylvania), Warren Schwartz

most systematic inquiry into the costs and benefits of judicial review. Its conclusions are a good place to start to answer this question.

Mashaw and his colleagues doubted the value of judicial review.²⁹ They compared the competence of generalist federal judges with specialist ALJs, suggesting that the latter's superior expertise called into question the corrective value of judicial review. Even if federal judges correct errors, they continued, the small number of cases appealed to the federal courts relative to the number of claims the agency decides undercut judicial review's ameliorative potential.³⁰ Also, this small number allows ALJs to remain indifferent to federal court review,³¹ and, for the most part, judicial opinions have little precedential value within the agency.³² Finally, the small number of appeals detracts from whatever "legitimizing function" judicial review entailed in the eyes of claimants.³³ Mashaw reiterated these lessons in his equally important 1983 study of disability claims adjudication.³⁴

Events since Mashaw's investigation may justify revisiting these conclusions. During the 1980s, the implementation of continuing disability review, a process by which existing beneficiaries are reexamined to determine their continuing entitlement to benefits, drew significant criticism, including from the federal courts.³⁵ The agency's treatment of claimants

(Georgetown University), Paul Verkuil (then University of North Carolina), and Milton Carrow (National Center for Administrative Justice). We refer only to Prof. Mashaw by name in the text for the sake of brevity and certainly not to slight the equal contributions of his distinguished co-authors.

²⁹ In his 1983 study of social security disability claims adjudication, Prof. Mashaw summarized these earlier findings as follows: "[M]y colleagues and I found that the tens of thousands of judicial review proceedings that have been held since the disability program's inception have either had no perceptible impact on its functioning or have made it worse." JERRY L. MASHAW, BUREAUCRATIC JUSTICE: MANAGING SOCIAL SECURITY DISABILITY CLAIMS 7 (1983).

³⁰ MASHAW ET AL., *supra* note 28 at 138-139; *id.* at 147.

³¹ *Id.* at 140.

³² *Id.* at 140-141.

³³ *Id.* at 147.

³⁴ MASHAW, BUREAUCRATIC JUSTICE, *supra* note 29 at 185-190.

³⁵ MARTHA DERTHICK, AGENCY UNDER STRESS: THE SOCIAL SECURITY ADMINISTRATION IN AMERICAN GOVERNMENT 36 (1990). Prof. Mashaw touched upon continuing disability review in his 1983 study and anticipated

with mental impairments sparked particularly intense resistance.³⁶ The episode “touched off a major battle between the agency and the courts,” as one prominent commentator observed, and the federal courts prevailed.³⁷ Both in individual cases and in class actions, the federal courts insisted on changes to disability claims adjudication, ones that have had long-lasting effect.³⁸

The influence of case law on agency decision-making has also changed. In 1978, Mashaw and his colleagues observed that the acquiescence process, or the process by which the agency accepts a particular circuit decision as binding for decision-makers within the circuit, had “lapsed into desuetude.”³⁹ Now more than four dozen acquiescence rulings have been issued,⁴⁰ and more opinions have influenced agency decision-making in equally important ways.⁴¹ Case law is responsible for the treating physician rule. It prompted rulemaking by the agency, and it remains one of the most significant doctrinal determinants of claim outcomes.⁴² The agency continues to discourage ALJs from relying on district court decisions or circuit decisions to which it has not acquiesced.⁴³ But, as discussed below, the ALJs we interviewed expressed more equivocal attitudes toward the instructive value of case law than what Mashaw and his colleagues documented in the 1970s.⁴⁴

We also question whether the argument against the corrective value of judicial review

some of the problems that it ultimately created. MASHAW, BUREAUCRATIC JUSTICE, *supra* note 29 at 174-175. But he completed his study before the process ran its course.

³⁶ JENNIFER L. ERKULWATER, DISABILITY RIGHTS AND THE AMERICAN SOCIAL SAFETY NET 110 (2006).

³⁷ DERTHICK, AGENCY UNDER STRESS, *supra* note 35 at 135 (1990).

³⁸ SOCIAL SECURITY ADMINISTRATION, HEARINGS, APPEALS, AND LITIGATION LAW MANUAL § I-5-4, at www.ssa.gov/OP_Home/hallex/hallex/html [hereinafter HALLEX] (table of contents identifying more than four dozen cases as affecting claims adjudication policy).

³⁹ MASHAW ET AL., *supra* note 28 at 110.

⁴⁰ https://ssa.gov/OP_Home/rulings/ar-toc.html.

⁴¹ The agency has published a large number of opinions as Social Security Rulings. See www.ssa.gov/OP_Home/rulings/rulfind3.html

⁴² ADMINISTRATIVE CONFERENCE OF THE UNITED STATES, SSA DISABILITY BENEFITS PROGRAMS: ASSESSING THE EFFICACY OF THE TREATING PHYSICIAN RULE, April 3, 2013, at 7-13.

⁴³ Soc. Sec. Ruling 96-1p; Memorandum to All Administrative Law Judges and All Senior Attorneys from Debra Bice, Chief Administrative Law Judge, Jan. 11, 2013, at 2 (on file with authors). Compare MASHAW ET AL., *supra* note 28 at 110-115 (describing the agency’s attitude toward case law in the 1970s).

⁴⁴ Mashaw and his colleagues reported that, “[i]n our interviews, ALJs and their legal assistants seemed remarkably unconcerned with the activities of their local reviewing courts.” MASHAW ET AL., *supra* note 28 at 140.

rooted in comparative institutional competence is as compelling as it might have been in 1978. Caseloads have fundamentally transformed ALJs' jobs. Based on their interviews, Mashaw and his colleagues estimated that an ALJ in 1978 averaged five hours on a case.⁴⁵ The ALJs we interviewed consistently reported a figure closer to two-and-a-half hours.⁴⁶ ALJs could afford to be "sharply divided" on the extent to which they delegated decision writing to staff attorneys in 1978.⁴⁷ Decision writing by staff attorneys and others is now an integral part of the process within most hearing offices. Any assessment of comparative institutional competence has to weigh the ALJ's superior expertise against the relative surfeit of time for decision-making that federal judges enjoy.

Judicial review undoubtedly has significant costs for the administration of our national disability claims system. The federal courts have a doubtful capacity to help engineer effective agency processes to be administered on a mass scale. Judicial insistence upon certain approaches to decision-making may create significant administrative burden with uncertain benefits.⁴⁸ The division of the federal courts into ninety-four districts and an inconstant commitment to uniformity in the elaboration of federal law by the twelve circuits create serious coordination difficulties for an agency committed to the consistent administration of a federal program.⁴⁹

But judicial review also has unquestionable benefits. The federal courts can act as a corrective to effects, whether intended or not, of what Mashaw calls "bureaucratic imperatives" –

⁴⁵ MASHAW ET AL., *supra* note 28 at 43.

⁴⁶ This figure approximates what we heard from most of the ALJs we interviewed. A report prepared for the Association of Administrative Law Judges in connection with litigation reported a similar figure. Cheryl Paullin et al., *Administrative Law Judge Work Analysis Study*, Nov. 12, 2015, at iii, at <https://drive.google.com/file/d/0B2kAAfgH45ClZjRGejdwSkwyNXc/view?pli=1> (last visited Mar. 31, 2016) (also on file with authors); *id.* at 42.

⁴⁷ MASHAW ET AL., *supra* note 28 at 90.

⁴⁸ MASHAW, BUREAUCRATIC JUSTICE, *supra* note 29 at 186-187; DERTHICK, AGENCY UNDER STRESS, *supra* note 35 at 131.

⁴⁹ DERTHICK, AGENCY UNDER STRESS, *supra* note 35 at 131.

aggregate priorities arrived at centrally by the agency that can change decision-making in a particular way.⁵⁰ For many claimants dogged enough to get to federal court, judicial review corrects errors that would otherwise deny benefits to those rightly entitled to them.

Despite calls for replacement dating back at least to Mashaw and his colleagues in 1978, Congress has left the present system of judicial review in place. Perhaps the inefficiencies and burdens that judicial review generates for the agency are not just a bug but part of the system's intended design. At any rate, we detected little support for a fundamental transformation in the present system of judicial review in the many interviews we conducted. Social security cases may often frustrate federal judges, but several stressed to us how highly they value their power to review ALJ decisions.⁵¹ Claimant representatives⁵² understandably place great value on access to the federal courts; as one told us, a federal court appeal for some claimants "is a matter of life and death."⁵³ Commentators have repeatedly proposed the creation of a specialized Article I court for the review of social security cases.⁵⁴ Almost no one we interviewed – federal judges, agency lawyers and officials, claimant representatives, or ALJs – favored this recommendation.

We assume for the purposes of this Report that the present system of judicial review will remain largely in place.⁵⁵ How to improve this system is the subject of the discussion that follows.

⁵⁰ MASHAW, BUREAUCRATIC JUSTICE, *supra* note 29 at 173. The best example is the resistance federal courts offered to continuing disability review in the 1980s. ERKULWATER, *supra* note 36 at 108 (2006).

⁵¹ *E.g.*, Federal Judge 11 (insisting that claimants should get Article III review); Federal Judge 19 (insisting that social security cases are important and should remain on district courts' dockets); Federal Judge 18 (doubting that specialization in appellate review would improve justice); Federal Judge 24 at 7 (insisting that specialization doesn't work well).

⁵² Readers should not confuse this term, which is commonly used in the manner that we do here, with "claims representatives." The latter work for the agency, where they help to review initial claims for benefits.

⁵³ Claimant Representative 1 at 6.

⁵⁴ *E.g.*, Verkuil & Lubbers, *supra* note 23 at 778.

⁵⁵ We consider and reject the possibility of replacing district court review with review by an Article I court in Part VI.

Part II. The Path of a Disability Claim

A number of the federal judges we interviewed candidly admitted that they know very little about how a disability claim works its way to their chambers.⁵⁶ The basics, of course, are familiar.⁵⁷ Claimants file an initial application, eventually appear at a hearing before an ALJ, and then seek Appeals Council review before heading to federal court. But a more detailed understanding of the path a claim follows is necessary for federal judges to understand fully what exactly they are reviewing and why it takes the form it does. This Part addresses gaps in judicial knowledge with an extensive description of this path. Two important determinants, caseload pressures and quality assurance initiatives, influence the composition of those cases that reach federal district courts. This Part discusses these factors as well, to deepen federal judicial knowledge of the influences that affect the quality of agency decisions they review.

Two provisos need mention. First, what follows does not cover every aspect of the claims adjudication process. We have selected details for inclusion based on what federal judges expressed particular interest in or confusion about in our conversations with them. Second, agency personnel in some of our interviews described their jobs in ways that may depart from agency policy. We have tried to indicate where this is so, and we emphasize that our interview data offer impressions but not rigorous proof of what happens inside the agency.

A. The Disability Claims Adjudication Process

1. The Initial Application

The quest for disability benefits begins with an application filed in one of SSA's 1,300

⁵⁶ *E.g.*, Federal Judge 8 at 3 (“I know nothing”); Federal Judge 9 at 4 (indicating no knowledge of processes within the agency); Federal Judge 10 at 3 (indicating knowing “next to nothing” about agency processes); Federal Judge 11 at 4 (reporting knowing “zero” about agency processes).

⁵⁷ For a thorough but concise summary of the process, see SOCIAL SECURITY ADVISORY BOARD, ASPECTS OF DISABILITY DECISION MAKING: DATA AND MATERIALS, Feb. 2012, at 87-88; *see also* Krent & Morris, *supra* note 1 at 373-377.

field offices, or online or by telephone.⁵⁸ A “claims representative,” the first SSA employee the applicant encounters, assembles some of the information necessary to determine whether a claimant is entitled to benefits.⁵⁹ The claims representative prepares a “disability report,” which includes information on the claimant’s work history, the alleged disability onset date, and the claimant’s medical providers.⁶⁰ The claims representative also creates a “certified electronic folder,”⁶¹ or a digital receptacle into which the claims representative, the claimant, her representative, and other personnel throughout the adjudication process deposit medical records and other claim-related information.⁶²

The claims representative verifies non-medical aspects of the claim.⁶³ She then forwards the claim to a “Disability Determination Service” (“DDS”), a state-run agency that operates under the SSA’s guidance and pursuant to federal law.⁶⁴ A “disability examiner” gathers as many of the claimant’s medical records as possible and then determines whether the claimant is disabled. In the majority of DDS offices, a medical or psychological consultant who works for the state agency may assist the disability examiner with this determination.⁶⁵ Some DDS offices have followed a “single decision maker” model that allows the disability examiner to decide some claims without a medical consultant’s signature.⁶⁶ In most states, the claimant can ask for

⁵⁸ ASPECTS OF DISABILITY DECISION MAKING, *supra* note 57 at 87.

⁵⁹ SOCIAL SECURITY ADMINISTRATION, PROGRAMS OPERATION MANUAL SYSTEM, § DI 10005.001, *at* www.ssa.gov/poms.nfs/lnx/0410005001 (last visited Nov. 11, 2015) [hereinafter “POMS”]. This person typically has a college degree but not medical training. *E.g.*, <https://www.socialsecurity.gov/ny/jobs-general.htm>; <http://www.dol.gov/oasam/doljobs/college-guide-qualify.htm>.

⁶⁰ POMS, *supra* note 59 § DI 10005.005.

⁶¹ *Id.*

⁶² *Id.* § DI 81001.005.

⁶³ *See* www.ssa.gov/disability/determination.htm (last visited Nov. 11, 2015); POMS, *supra* note 59 § DI 10005.001B.

⁶⁴ ASPECTS OF DISABILITY DECISION MAKING, *supra* note 57 at 87.

⁶⁵ *E.g.*, www.disabilitysecrets.com/resources/disability/find-out-what-type-medical-consultant-reviewed.

⁶⁶ *See generally* 20 C.F.R. §§ 404.906 and 416.1406; OFFICE OF THE INSPECTOR GENERAL, SOCIAL SECURITY ADMINISTRATION, AUDIT REPORT, SINGLE DECISIONMAKER MODEL – AUTHORITY TO MAKE CERTAIN DISABILITY DETERMINATIONS WITHOUT A MEDICAL CONSULTANT’S SIGNATURE, Aug. 2013, *at*

reconsideration if the DDS office denies the claim.⁶⁷ A different pair of examiners and consultants decides these requests.⁶⁸

2. The Hearing Office

Any discussion of adjudication at the ALJ level must begin with a sobering reality: the 1300 or so ALJs working within more than 160 hearing offices nationwide⁶⁹ are each asked to render 500-700 “legally sufficient” dispositions per year.⁷⁰ The agency set this goal in 2007 in response to concerns about an immense backlog and lengthy delays claimants had to endure,⁷¹ and concomitant political pressure to clear it.⁷² At the time, the majority of ALJs issued fewer than 500 decisions annually.⁷³ By 2011, two-thirds of ALJs surpassed the 500 case threshold.⁷⁴ In FY 2013, the average ALJ disposed of 48 cases per month,⁷⁵ although ALJ disposition rates

<http://oig.ssa.gov/sites/default/files/audit/full/pdf/A-01-12-11218.pdf>. The Bipartisan Budget Act of 2015 will eliminate the agency’s single decision maker authority effective November 2, 2016. *See* Pub. L. No. 114-74, § 832.

⁶⁷ In so-called “prototype” states claimants cannot request reconsideration. POMS, *supra* note 59 § DI 12015.100. These include New Hampshire, parts of New York, Pennsylvania, Alabama, Michigan, Louisiana, Missouri, Colorado, parts of California, and Alaska. *Id.*; *see also* HALLEX, *supra* note 38 § I-2-4-98. The evidence is mixed, but in general claimants do not fare differently at the ALJ level whether they come from a prototype jurisdiction or not. HAROLD J. KRENT & SCOTT MORRIS, STATISTICAL APPENDIX ON ACHIEVING GREATER CONSISTENCY IN SOCIAL SECURITY DISABILITY ADJUDICATION: AN EMPIRICAL STUDY AND SUGGESTED REFORMS, April 2013, at 50-51 (April 2013); ASPECTS OF DISABILITY DECISION MAKING, *supra* note 57 at 10-11.

⁶⁸ ASPECTS OF DISABILITY DECISION MAKING, *supra* note 57 at 88.

⁶⁹ SOCIAL SECURITY ADMINISTRATION, ANNUAL STATISTICAL SUPPLEMENT TO THE SOCIAL SECURITY BULLETIN, 2014, at Table 2.F8; *id.* Table 2.F1; *see also* OFFICE OF THE INSPECTOR GENERAL, SOCIAL SECURITY ADMINISTRATION, AUDIT REPORT, THE SOCIAL SECURITY ADMINISTRATION’S EFFORTS TO ELIMINATE THE HEARING BACKLOG, A-12-15-15005, Sept. 2015, at 5 (showing fluctuations in ALJ numbers). The Fall 2015 budget deal provides for a significant growth in the size of the ALJ corps, to over 1900 ALJs, by 2017. *E.g.*, Mark Schleifstein, *Social Security Disability Pressure Eased by Budget Compromise*, NEW ORLEANS TIMES-PICAYUNE, Oct. 30, 2015.

⁷⁰ Letter from Chief Administrative Law Judge Frank A. Cristaudo to Colleagues, Oct. 31, 2007, at 2, at http://www.aalj.org/system/files/documents/exhibits_a-f_usdc.pdf.

⁷¹ *Id.*

⁷² *E.g.*, *Social Security Disability Backlogs*, Hearing Before the Subcommittee on Social Security of the Committee on Ways and Means, U.S. House of Representatives, 110th Cong., 1st Sess., Feb. 14, 2007; David Harrison, *Social Security: Short-Term Progress, Long-Term Perils*, CONG. QUART. WEEKLY, Feb. 2, 2013.

⁷³ Association of Administrative Law Judges v. Colvin, 777 F.3d 402, 403 (7th Cir. 2015).

⁷⁴ KRENT & MORRIS, STATISTICAL APPENDIX, *supra* note 67 at 7. Between 2007 and 2011 the agency hired hundreds of new ALJs. *See* Statement of Michael J. Astrue, Commissioner, Social Security Administration, Before the House Committee on Ways and Means, Subcommittee on Social Security and the House Committee on the Judiciary Subcommittee on the Courts, Commercial and Administrative Law, July 11, 2011; *see also* *Leading the Hearing and Appeals Process Into the Future: A Plan for Compassionate And Responsive Service*, Jan. 13, 2016, at 5, at <http://www.scribd.com/doc/296323512/Cares-Plan-For-Dealing-With-ODAR-Backlog> (last visited Mar. 29, 2016).

⁷⁵ 2014 ANNUAL STATISTICAL SUPPLEMENT, *supra* note 69 Table 2.F8.

varied considerably.⁷⁶ This number has come down slightly in recent years, in part due to an increased agency emphasis on policy compliant decision-making.⁷⁷

a. Pre-Hearing Workup and Hearing

Claimants, three-fourths of whom are represented at the hearing stage,⁷⁸ must file requests for hearings within sixty days of the final denial of claims at the DDS level.⁷⁹ After an initial case analysis, cases may be screened by attorney adjudicators or undergo other processing functions. Hearing office staff members are then assigned cases to review, analyze, and prepare for a hearing. As they do so, they organize the medical evidence already in the electronic folder, add newly submitted evidence to the right sections of the electronic folder, and organize earnings records and other non-medical information.⁸⁰ During this process ALJs get assigned to cases on a rotational basis. Some ALJs have standing orders that instruct hearing office staff to acquire information that is missing, such as workers' compensation records, or to schedule vocational expert or medical expert testimony if a case warrants it.⁸¹ Other ALJs will take a first look at a case months in advance of a hearing date, to determine what sort of workup it requires, and then give hearing office staff more tailored instruction for the pre-hearing workup.⁸²

The inquisitorial, non-adversarial nature of the claims adjudication process obliges an ALJ to “develop the record.”⁸³ Federal judges commonly fault ALJs for their failures to discharge this duty adequately, although expectations about what ALJs can plausibly accomplish may be unreasonable in some instances. An example involves the enforcement of subpoenas

⁷⁶ KRENT & MORRIS, STATISTICAL APPENDIX, *supra* note 67 at 5-6. In one of the years covered by the Krent and Morris study, one ALJ decided 3,620 claims. *Id.*

⁷⁷ OIG, EFFORTS TO ELIMINATE THE HEARING BACKLOG, *supra* note 69 at 4.

⁷⁸ KRENT & MORRIS, STATISTICAL APPENDIX, *supra* note 67 at 48.

⁷⁹ HALLEX, *supra* note 38 § I-2-0-50.

⁸⁰ Agency Official 1 at 1.

⁸¹ ALJ 5 at 3; ALJ 3 at 3.

⁸² ALJ 6 at 1; ALJ 7 at 1.

⁸³ 20 C.F.R. §§ 404.1512(d)-(e), 416.912(d)-(e).

ALJs issue to medical providers for relevant records.⁸⁴ To some federal courts, especially in *pro se* cases, the mere issuance of a subpoena does not discharge the ALJ's obligation to develop the record when the person or entity being subpoenaed does not respond.⁸⁵ An ALJ who seeks a subpoena's enforcement, however, must trigger a cumbersome process.⁸⁶ It can require the Office of General Counsel to work with the United States Attorney to seek enforcement of the subpoena, and ultimately federal court involvement.⁸⁷

Most of the twenty-four ALJs we interviewed subscribe to what one labeled a "just in time" approach to case review.⁸⁸ An ALJ using this method first looks at a case anywhere from one day to a week before the hearing.⁸⁹ One ALJ acknowledged to us that an earlier review would be better; the ALJ, for example, could make sure that she had solicited the right expert testimony given the alleged impairments and the makeup of the medical records.⁹⁰ A variety of factors, however, can compel the just in time approach, including an issue of significant concern to ALJs. Claimant representatives often submit voluminous sets of medical records up to and even after the hearing date.⁹¹ One ALJ told us that late-arriving evidence surfaces in 90% of cases;⁹² others said this happens "every day."⁹³ Explanations for the phenomenon vary.⁹⁴

⁸⁴ HALLEX, *supra* note 38 § I-2-5-78; 20 C.F.R. §§ 404.950(d)(1) and 416.1450(d)(1).

⁸⁵ *E.g.*, *Sanchez v. Barnhart*, 329 F. Supp. 2d 445, 451 (S.D.N.Y. 2004); *Suriel v. Comm'r of Social Security*, Civ. No. 05-1218, 2006 WL 2516429, at *6 (E.D.N.Y. Aug. 29, 2006); *Brandow v. Comm'r of Social Security*, Civ. No. 05-917, 2009 WL 2971543, at *5 n.6 (N.D.N.Y. Sept. 11, 2009). *But see* *Serrano v. Barnhart*, Civ. No. 02-6372, 2005 WL 3018256, at *3-4 (S.D.N.Y. Nov. 1, 2005); *Friedman v. Astrue*, Civ. No. 07-3651, 2008 WL 3861211, at *4-5 (S.D.N.Y. Aug. 19, 2008).

⁸⁶ HALLEX, *supra* note 38 § I-2-5-82.

⁸⁷ 42 U.S.C. § 405(e). *Cf.* *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).

⁸⁸ ALJ 8 at 3.

⁸⁹ *E.g.*, ALJ 9 at 1 (one week); ALJ 10 at 1 (one week); ALJ 11 at 1 (2-3 days before); ALJ 12 at 1 (one week); ALJ 7 at 1 (10 days); ALJ 8 (one day before); ALJ 1 at 1 (one day); ALJ 13 at 1 (one week); ALJ 14 at 1 (one day); ALJ 2 at 1 (a couple of days).

⁹⁰ ALJ 7 at 1.

⁹¹ *See generally* ADMINISTRATIVE CONFERENCE OF THE UNITED STATES, FINAL REPORT, SSA DISABILITY BENEFITS ADJUDICATION PROCESS: ASSESSING THE IMPACT OF THE REGION I PILOT PROGRAM, Dec. 2013, at 10; *id.* at 29.

⁹² ALJ 6 at 2; *see also* ALJ 14 at 1 (70-90% of cases).

⁹³ ALJ 13 at 1; *see also* ALJ 12 at 1 (insisting that this practice is the rule, not the exception).

Whatever the reason for the problem, late-arriving records can turn an earlier case review into a fools' errand, several ALJs insisted to us. If four hundred pages of medical records show up a day before the hearing, their arrival undermines much of the preparation an ALJ may have done three weeks earlier.

We asked all of the ALJs we interviewed to estimate how much time they spend on various tasks. Several described spending a day preparing a day's worth of hearings, usually six or eight.⁹⁵ ALJs are expected to schedule 10-20 hearings per week.⁹⁶ A union contract does not permit ALJs to work overtime, so this schedule has implications for the amount of time ALJs have for each case.⁹⁷ ALJs we interviewed reported spending about an hour or so reviewing claimants' records and preparing for a hearing,⁹⁸ a figure consistent with a previous study.⁹⁹ Medical records routinely total hundreds of pages in length.¹⁰⁰ Some ALJs told us that they try

⁹⁴ See generally Frank S. Bloch et al., *Developing a Full and Fair Evidentiary Record in a Nonadversary Setting: Two Proposals for Improving Social Security Disability Adjudications*, 25 CARDOZO L. REV. 1, 8 (2003) (discussing possible explanations). Some ALJs blamed this phenomenon on cynical or negligent claimant representatives. ALJ 13 at 2; ALJ 2 at 1. Claimant representatives stressed the difficulty and expense of obtaining records from medical providers. Claimant Representative 2 at 3; Claimant Representative 3 at 1.

⁹⁵ E.g., ALJ 15 at 2 (eight hearings per day); ALJ 10 at 2 (six hearings per day, one day to prepare one day's worth of hearings); ALJ 16 (five hearings per day for a Hearing Office Chief Administrative Judge ("HOCALJ"), one day to prepare one day's worth of hearings).

⁹⁶ OFFICE OF THE INSPECTOR GENERAL, SOCIAL SECURITY ADMINISTRATION, CONGRESSIONAL RESPONSE REPORT, ADMINISTRATIVE LAW JUDGE AND HEARING OFFICE PERFORMANCE, A-07-08-28094, Aug. 2008, at 10. ALJs may not have hearings every week. Some ALJs we interviewed took a week to prepare a week's worth of hearings. Others presumably have hearings on some days and use other days to prepare for hearings and prepare decisions. In 2014, the agency issued a memorandum to ALJs providing that "scheduling an average of at least fifty (50) cases for hearing per month will generally signify a reasonably attainable number" Memorandum to All Administrative Law Judges from Debra Bice, Chief Administrative Law Judge, Feb. 18, 2014, at 2 (on file with authors).

⁹⁷ A full work-year for ALJs includes 2,087 hours. For a detailed description of how many hours ALJs actually spend on cases, see generally Paullin et al., *supra* note 46 at iii. This study reports that ALJs spend nearly forty hours per month on tasks not associated with deciding cases. *Id.* at 39. ALJs may work credit hours between 6:30 AM and 6:00 PM on non-regular work days.

⁹⁸ E.g., ALJ 9 at 4 (30-60 minutes); ALJ 15 (two hours); ALJ 17 at 2 (45-60 minutes); ALJ 18 at 3 (one hour); ALJ 19 (15 minutes to three hours, depending on the case's complexity).

⁹⁹ OIG, ADMINISTRATIVE LAW JUDGE AND HEARING OFFICE PERFORMANCE, *supra* note 96 at 9 (observing that higher producing ALJs spend less than an hour reviewing cases before a hearing).

¹⁰⁰ In FY 2014, the average case file included 652 pages, including 367 pages of medical records. Paullin et al., *supra* note 46 at 7.

to review every page of a file.¹⁰¹ To others, however, reading every page is an “impossible” undertaking.¹⁰² One ALJ conceded that he “just hits the high notes” in his pre-hearing review and thus that he “get[s] surprised” at the hearing.¹⁰³ This approach does not comply with agency policy, but this ALJ was not the only one to confess to it.¹⁰⁴

A hearing typically lasts anywhere from fifteen minutes to an hour.¹⁰⁵ Several ALJs told us that they do their best to hold the hearing, even if a claimant representative produces a stack of medical records as the hearing begins.¹⁰⁶ One ALJ described this practice as the agency’s preference,¹⁰⁷ although the formal policy is otherwise. About a quarter of hearings proceed by video,¹⁰⁸ enabling the agency to shift cases from particularly burdened offices to those in different locations that might have more capacity.¹⁰⁹ Some of these video hearings proceed before ALJs in five National Hearing Centers, which opened starting in 2007 expressly to receive cases transferred from “heavily backlogged” hearing offices nationwide.¹¹⁰

b. Writing the Decision

¹⁰¹ *E.g.*, ALJ 3 at 1. This ALJ acknowledged a lower than average disposition rate. *Id.* at 1 (suggesting that most ALJs write more decisions).

¹⁰² ALJ 14 at 1; *see also* ALJ 18 at 3; ALJ 20 at 5.

¹⁰³ ALJ 21 at 3.

¹⁰⁴ ALJ 19 at 1

¹⁰⁵ ALJ 17 at 2. Bloch et al., *supra* note 94 at 26 (2003); OIG, ADMINISTRATIVE LAW JUDGE AND HEARING OFFICE PERFORMANCE, *supra* note 96 at 16.

¹⁰⁶ ALJ 13 at 1; ALJ 14 at 2; ALJ 11 at 2.

¹⁰⁷ ALJ 16 at 3.

¹⁰⁸ Between September 26, 2015, and February 26, 2016, 169 hearing offices reported conducting 59,560 hearings by video and 226,572 hearings total. *See* https://www.ssa.gov/appeals/DataSets/06_Hearings_Held_InPerson_Video_Report.html. This figure represents a significant increase over the past decade. *See* OFFICE OF THE INSPECTOR GENERAL, SOCIAL SECURITY ADMINISTRATION, AUDIT REPORT: USE OF VIDEO HEARINGS TO REDUCE THE HEARING CASE BACKLOG, A-05-08-18070, Apr. 2011, at 2 (reporting figures from 2009 and earlier).

¹⁰⁹ *Id.* at 4.

¹¹⁰ OFFICE OF THE INSPECTOR GENERAL, SOCIAL SECURITY ADMINISTRATION, AUDIT REPORT: THE ROLE OF NATIONAL HEARING CENTERS IN REDUCING THE HEARINGS BACKLOG, A-12-11-11147, Apr. 2012, at 1-2 (reporting figures from 2009 and earlier). In FY 2011, National Hearing Centers accounted for 4.2% of ALJ dispositions. *Compare id.* at 3 (reporting 33,324 NHC dispositions), *with* <https://www.ssa.gov/policy/docs/statcomps/supplement/2013/2f8-2f11.html#table2.f9> (reporting 793,563 ALJ dispositions). Between Between September 26, 2015, and February 26, 2016, National Hearing Centers conducted 7,650 hearings, out of a total of 226,572 hearings conducted by 169 offices, or 3.4%. *See* https://www.ssa.gov/appeals/DataSets/06_Hearings_Held_InPerson_Video_Report.html.

ALJs do not necessarily decide claims right after the hearing. Newly submitted records might prompt the ALJ to request a consultative exam, which may entitle the claimant to a supplemental hearing.¹¹¹ If testimony elicited at the hearing or late arriving records affect the ALJ's understanding of the claimant's impairments, the ALJ might have to send interrogatories to a medical expert to assist in determining the claimant's residual functional capacity ("RFC").¹¹² Claimant representatives might provide yet more records a month later. These post-hearing developments are irritants, one ALJ told us, because they might require the ALJ to prepare the case again.¹¹³ More often, however, ALJs make decisions immediately after hearings, when memories of testimony and the documentary evidence are sharp.¹¹⁴ The key task is to prepare instructions for decision writers, who most often draft the ALJ's decision.

Decision writers play a central role in the claims adjudication process but are all but unknown to federal judges.¹¹⁵ This unawareness is striking, akin to an appellate judge not knowing that district judges have clerks or what clerks do. A federal judge cannot possibly put ALJs' decisions in proper institutional context unless she appreciates the decision writer's role.

Decision writers are typically either attorneys or "paralegal specialists," the latter of whom tend to come up from within SSA's ranks.¹¹⁶ They often work alongside ALJs in hearing offices, although in recent years SSA has operated two "national case assistance centers," or centralized offices where decision writers draft decisions for ALJs around the country.¹¹⁷ The SSA has set a target of three decision writers for every two ALJs, but most hearing offices fall a

¹¹¹ HALLEX, *supra* note 38 § I-2-7-30.

¹¹² HALLEX, *supra* note 38 § I-2-5-45.

¹¹³ One ALJ insisted he could decide an additional 100 cases a year "easily" if he didn't have to perform this sort of post-hearing cleanup. ALJ 19 at 4.

¹¹⁴ *E.g.*, ALJ 13 at 2; ALJ 10 at 1-2.

¹¹⁵ Federal Judge 12 at 1 (expressing uncertainty if ALJs have "clerks"); Federal Judge 6 at 4 (mentioning that only recently after fourteen years discovered that ALJs have decision writers).

¹¹⁶ One decision writer we interviewed, for example, started her career with SSA as a claims representative in a field office. Decision Writer 1 at 1.

¹¹⁷ See www.socialsecurity.gov/org/orgdcdr.htm.

good deal short of this goal.¹¹⁸ Decision writers are not usually assigned to particular ALJs but instead work in a pool,¹¹⁹ a scheme many ALJs and decision writers dislike.¹²⁰ Within these pools, decision writers may be organized differently from hearing office to hearing office. A decade ago, for example, decision writers in one hearing office were given particular topics, such drug and alcohol cases, to become expert on, and difficult cases involving these topics would get assigned accordingly.¹²¹ More commonly, decisions are put in a queue, to be assigned to the next decision writer available. Decision writers are expected to spend roughly four hours on a decision granting benefits and eight hours on a decision denying them, although the SSA has recently crafted more nuanced expectations for decision writer productivity.¹²² One decision writer told us that “two-dayers are pretty rare,”¹²³ and another described 20-30 decisions per month as an acceptable pace.¹²⁴

Decision writers do not typically attend hearings, although they may listen to the hearing recording, and they base their drafts on written instructions from ALJs. These instructions vary significantly. Although some ALJs remain committed to their own forms, typing into them or filling them out by hand,¹²⁵ the SSA has pushed for standardization in recent years. The “electronic benchbook” (“eBB”), a web-based application that aids in documenting, analyzing,

¹¹⁸ OFFICE OF THE INSPECTOR GENERAL, SOCIAL SECURITY ADMINISTRATION, AUDIT REPORT: HEARING OFFICE PERFORMANCE AND STAFFING, A-12-08-28088, Feb. 2010, at 7; *see also* Agency Official 1 at 3 (indicating that the ratio is typically 1:1). One decision writer we interviewed insisted that the ratio should be 2-3 decision writers per ALJ. Decision Writer 1 at 1. This is the ratio that prevails in National Hearing Centers. OIG, THE ROLE OF NATIONAL HEARING CENTERS, *supra* note 110, at 6.

¹¹⁹ Decision Writer 2 at 1. Decision writers are assigned differently in the National Hearing Centers. Because the National Hearing Centers render only a small percentage of the total number of ALJ dispositions, however, we believe the description is generally applicable.

¹²⁰ Jeffrey S. Wolfe, *Civil Justice Reform in Social Security Adjudications*, 33 J. NAT’L ASS’N ADMIN. L. JUDICIARY 137, 176 (2013) (reporting results of a survey of ALJs); Decision Writer 2 at 1 (indicating from a decision writer’s point of view that the assignment decision writers to individual ALJs would be better, provided that the ALJ weren’t difficult to work with).

¹²¹ Decision Writer 2 at 2. This practice ended a decade ago.

¹²² *E.g.*, Decision Writer 3 at 3; Decision Writer 4 at 2; Decision Writer 5 at 2.

¹²³ Decision Writer 3 at 3. Another decision writer told us that he “hardly ever” takes more than eight hours to draft a decision. Decision Writer 6 at 1.

¹²⁴ Decision Writer 7 at 2.

¹²⁵ ALJ 9 at 2; ALJ 15 at 2; ALJ 19 at 1.

and adjudicating cases, generates instructions for some ALJs. More common are instructions based on the Findings Integrated Template (“FIT”), a tool ALJs began using in 2006 that generates templates for various types of disability decisions.¹²⁶

The SSA has pushed ALJs to give more than a thumbs up/thumbs down in their instructions; at a minimum, an adequate set should include information reflecting the ALJ’s determination for each step of the five-step sequential process for evaluating a claim.¹²⁷ Several of the decision writers we interviewed, however, reported that instructions routinely omit basic findings, such as whether the claimant has a severe impairment¹²⁸ or what the claimant’s residual functional capacity is.¹²⁹ ALJs, some writers told us, often do not address the weight they assign to treating physician opinions, and often they do not include any information in their instructions to explain a “symptom evaluation” finding.¹³⁰ (The agency now uses the term “symptom evaluation” in lieu of “credibility.”)¹³¹ One ALJ told us that, if he were to draft instructions in

¹²⁶ See generally OFFICE OF THE INSPECTOR GENERAL, SOCIAL SECURITY ADMINISTRATION, OFFICE OF DISABILITY ADJUDICATION AND REVIEW DECISION-WRITING PROCESS, A-2-09-19068, Nov. 2010, at 2.

¹²⁷ *Id.* at 2; see also Memorandum to All Administrative Law Judges, Attorney Adjudicators, and Decision Writers, from Debra Bice, Chief Administrative Law Judge, Feb. 27, 2012 (on file with authors) (describing adequate instructions); Memorandum to All Administrative Law Judges, from Debra Bice, Chief Administrative Law Judge, July 10, 2013 (on file with authors) (describing adequate instructions).

¹²⁸ Decision Writer 8 at 4.

¹²⁹ Decision Writer 7 at 1; see also Decision Writer 12 at 4 (explaining that “we have to reinvent the wheel” to draft a complete decision); Decision Writer 4 at 2 (explaining that “in quite a few instances” he has had to mine the record to understand why an ALJ wanted to rule the way the ALJ did). A claimant’s “residual functional capacity,” or RFC, represents the most that he or she can do despite his or her impairments. See 20 C.F.R. §§ 404.1545 and 416.945.

¹³⁰ One decision writer estimated that one out of twelve sets of instructions will include information to support a credibility finding, and two to three out of twelve will assign weight to treating physician opinions. Decision Writer 9 at 3. Another told us that in a “majority” of instances she has to figure out on her own what weight to give treating physician opinions. Decision Writer 3 at 2; see also Decision Writer 7 at 1 (explaining that “ideally” instructions would discuss “symptom evaluation” and weight given to treating physicians’ opinions, but more often than not they do not); Decision Writer 6 at 1 (explaining that “better” instructions address treating physician opinions); Decision Writer 4 at 1 (describing instructions that address weight given to physician evidence as at a “far extreme”). *But see* Decision Writer 10 at 2 (explaining that instructions “more or less” address symptom evaluation and treating opinion evidence).

¹³¹ Soc. Sec. Ruling 16-3p.

accordance with the agency’s guidelines, he would decide three cases per week.¹³² Some decision writers actually prefer this brevity and the responsibility that shifts to their shoulders.¹³³

The decision writers we interviewed take various approaches to drafting decisions. A couple reported that they go through all the records,¹³⁴ with an understanding that they act as a “backstop” to the ALJ.¹³⁵ Others “brows[e] through” medical records when they are too voluminous to examine page-by-page.¹³⁶ The SSA has encouraged decision writers to discuss concerns with ALJs when instructions have not addressed important issues. The decision writers with whom we spoke have mostly constructive relationships with ALJs and reported easy, collegial interactions. But ALJs can sometimes bristle at decision writers’ questions, one writer said,¹³⁷ and another told us that consulting with ALJs can slow things down.¹³⁸ Contrary to policy, several decision writers described routinely determining issues they felt a decision needed without guidance from an ALJ’s instructions.¹³⁹ Concerns that might change the claim’s outcome, however, typically get reported to ALJs before a draft’s completion,¹⁴⁰ and ultimate responsibility for all findings always remains with the ALJ.

¹³² ALJ 16 at 3. A report commissioned by the Association of Administrative Law Judges concluded that a decision crafted in a manner consistent with agency policy, including that regarding appropriate instructions, would require between five-and-a-half and eight-and-a-half hours of ALJ time, depending on case complexity. Paullin et al., *supra* note 46 at iii.

¹³³ “Just say ‘deny at Step Four’ and let me figure out why,” as one decision writer declared. Decision Writer 2 at 3; *see also* Decision Writer 11 at 1 (detailed instructions are “frustrating,” since a decision writer has to go through the record anyway).

¹³⁴ Decision Writer 1 at 3; Decision Writer 9 at 2 (explaining that she goes through the record “very thoroughly”).

¹³⁵ Decision Writer 1 at 4; Decision Writer 12 at 1.

¹³⁶ Decision Writer 13 at 1; *see also* Decision Writer 7 at 3 (explaining that he does not go through every page and instead skims records).

¹³⁷ *E.g.*, Decision Writer 2 at 4.

¹³⁸ Decision Writer 12 at 5.

¹³⁹ Decision Writer 13 at 4 (decision writers make their own determinations in 99.9% of cases); Decision Writer 8 at 4 (100% of cases); Decision Writer 1 at 3 (responding “oh yes” when asked whether decision writers find issues ALJs didn’t address in their instructions); Decision Writer 2 at 4 (estimating that she catches issues “20-30% of the time at least”); Decision Writer 4 at 2 (indicating that this happens “with some regularity”). *But cf.* Decision Writer 7 at 2 (suggesting that serious issues with ALJ’s decision comes up “very rarely”); Decision Writer 6 at 2 (indicating that an issue the ALJ didn’t address arises once every 2-3 weeks); Decision Writer 12 at 5.

¹⁴⁰ Decision Writer 8 at 4.

Decision writers do not write on a blank slate but instead insert particularized text providing the decision’s rationale into boilerplate that the FIT template generates. Ninety-eight percent of opinions are written on the FIT template, which is intended to ensure that all decisions meet a certain threshold of quality and completeness. For this reason, almost every ALJ decision will contain at least some generic language.¹⁴¹ The draft then goes back to the ALJ for review. Estimates of time spent reviewing drafts ranged from fifteen minutes¹⁴² to four hours;¹⁴³ thirty minutes was a common estimate.¹⁴⁴ More often than not ALJs do not revise drafts extensively.¹⁴⁵

3. The Appeals Council

Although every disability appeal passes through the Appeals Council before arriving at a district court, administrative records reveal little about what goes on within the SSA’s appellate body. The Appeals Council decides whether to grant review of a hearing decision, and if review is denied, then the claimant can proceed to federal court. These denial actions are documented in a brief order. Appeals Council processes, then, are particularly opaque to federal courts.¹⁴⁶

¹⁴¹ OIG, EFFORTS TO ELIMINATE THE HEARING BACKLOG, *supra* note 69 at 9.

¹⁴² ALJ 9 at 4.

¹⁴³ ALJ 11 at 4 (estimating spending four hours reviewing a weak writer’s drafts and twenty minutes reviewing a strong writer’s drafts).

¹⁴⁴ ALJ 15 at 3; ALJ 18 at 4. *See also* Paullin et al., *supra* note 46 at 48 (reporting that 77% of draft decisions require no more than an hour to review).

¹⁴⁵ Letter from Frank A. Cristaudo, Chief Administrative Law Judge, to Colleagues, Dec. 19, 2007, at 2 (on file with authors) (“We . . . ask the attorneys and paralegals to provide draft decisions that require little or no editing. We do not want the ALJ, or anyone else, to waste precious time re-drafting decisions.”). One ALJ told us that her “greatest frustration” is having to put her name on a decision that will get “butchered” on appeal because it is “so mediocre or poor.” But, because she has no more than “minimal time” to spend reviewing decisions, this ALJ has accepted such frustration as an inevitable by-product of her caseload. ALJ 20 at 3; *see also* ALJ 16 at 4 (acknowledging that he reviews drafts quickly). *See generally* OIG, ADMINISTRATIVE LAW JUDGE AND HEARING OFFICE PERFORMANCE, *supra* note 96 at 13 (describing estimates of how intensive of a review ALJs give decision writer drafts); Paullin et al., *supra* note 46 at 48 (observing that 23 out of 31 ALJs surveyed reported that 50% of draft decisions require little or no editing).

¹⁴⁶ For a detailed introduction to processes in the Appeals Council, see OFFICE OF THE INSPECTOR GENERAL, SOCIAL SECURITY ADMINISTRATION, AUDIT REPORT: REQUEST FOR REVIEW WORKLOADS AT THE APPEALS COUNCIL, A-12-13-13039, Mar. 2014.

The roughly 180,000 appeals claimants take annually from ALJ decisions get divvied up among the five “Program Review Divisions” within the Appeals Council. Typically, each division handles cases originating in states that fall within a maximum of two judicial circuits.¹⁴⁷ These divisions in turn are divided into branches, with each branch staffed by about 15-18 analysts.¹⁴⁸ 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.¹⁴⁹ Analysts¹⁵⁰ use a case-processing tool called the “Appeals Case Analysis Tool,” or ACAT, to deconstruct the ALJ’s decision and determine if it is policy compliant.¹⁵¹ ACAT asks analysts a series of questions about each of the sequential steps to determine if the Appeals Council should grant review of the ALJ’s decision.¹⁵² Some claimants file short briefs, typically no more than three pages or so, identifying alleged errors in the ALJ’s decision.¹⁵³ A brief can be little more than a pre-printed form letter that rattles off a laundry list of alleged errors.¹⁵⁴

The ACAT tool requires the analyst to address key issues in a case, and it prompts an analyst to justify a recommendation. Usually, analysts prepare “action documents” that include recommended actions along with the analysis generated by the ACAT tool.¹⁵⁵ According to one agency official’s description, these action documents typically range from a half of a page to a

¹⁴⁷ Appeals Council Organizational Chart, Sept. 26, 2011 (on file with authors).

¹⁴⁸ Agency Official 2, First Interview at 1.

¹⁴⁹ Agency Official 2, Second Interview at 1.

¹⁵⁰ The majority of analysts are attorneys. Some are paralegal specialists. Agency Official 3, First Interview at 1.

¹⁵¹ Agency Official 2, First Interview at 1; Agency Official 3, Third Interview at 1; OIG, REQUEST FOR REVIEW WORKLOADS, *supra* note 146 at E-1. ACAT is part of the “Appeals Review Processing System,” or ARPS. ARPS is described as follows:

ARPS can generate detailed and structured management information reports on receipts, adjusted receipts, dispositions, average processing time, pending, and average age of pending. It can sort reports by fiscal year, month, and week. The system can also produce listings based on a large selection of criteria, such as by projected age of the case, by case status, by the assigned analyst, by branch, type of case, and State. The system also allows the user to drill down to the case details.

Id. at E-1.

¹⁵² Agency Official 2, First Interview at 1.

¹⁵³ Agency Official 2, Second Interview at 2.

¹⁵⁴ Agency Official 2, Second Interview at 2.

¹⁵⁵ HALLEX, *supra* note 38 § I-2-3-10.

page-and-a-half of individualized text;¹⁵⁶ the ACAT tool produces a “one-page Facts of Finding memorandum.”¹⁵⁷ If an analyst recommends that the AC deny review, the case will go to either an Appeals Officer (“AO”) for or an Administrative Appeals Judge (“AAJ”). A recommendation that the AC grant review and remand will go to two AAJs. These officials will then formally decide the appeal.¹⁵⁸

Caseload pressures are immense at the Appeals Council. After an initial training period,¹⁵⁹ analysts are expected on average to prepare action documents for two cases each day.¹⁶⁰ Depending on complexity, AOs and AAJs decide anywhere from five to twelve cases per day.¹⁶¹ An analyst we interviewed told us that some analysts claim to go through every page in a case file, but that he did not think it feasible to do so.¹⁶² AOs and AAJs face the same time constraints.¹⁶³ Still, an experienced analyst, having reviewed hundreds of cases, can “flip through” medical records quickly to find a physician’s evaluative opinion, one analyst told us.¹⁶⁴ Extensive experience also enables AOs and AAJs to spot problems in ALJ decisions quickly.¹⁶⁵

At present Appeals Council adjudication mostly involves error correction. Unlike the Board of Immigration Appeals, another agency appellate body handling a large volume of cases, the Appeals Council rarely issues opinions or otherwise steers the elaboration of governing law

¹⁵⁶ Agency Official 2, First Interview at 1.

¹⁵⁷ OIG, REQUEST FOR REVIEW WORKLOADS, *supra* note 146 at E-1.

¹⁵⁸ OIG, REQUEST FOR REVIEW WORKLOADS, *supra* note 146 at 1.

¹⁵⁹ Agency Official 3, Third Interview at 2.

¹⁶⁰ Agency Official 2, First Interview at 2.

¹⁶¹ Agency Official 4 at 2. In FY 2012, the median AAJ decided 1,283 dispositions. The median AO rendered 2,049 dispositions. OIG, REQUEST FOR REVIEW WORKLOADS, *supra* note 146 at 10, 12.

¹⁶² Agency Official 5 at 2.

¹⁶³ Several claimant representatives relayed an impression to us that Appeals Council review is cursory. Claimant Representative 4 at 3 (“They have so very little time.”); Claimant Representative 5 at 2 (noting that the AC has “an incredibly limited amount of time”). *Compare* Agency Official 2, First Interview at 2 (insisting that some AOs and AAJs review medical records comprehensively).

¹⁶⁴ Agency Official 5 at 3.

¹⁶⁵ Agency Official 3, Third Interview at 2. An analyst told us that some AAJs and AOs delve deeply into records. Agency Official 2, Second Interview at 2.

through precedent.¹⁶⁶ Its personnel do, however, rely on their review of cases to identify issues for policy making through other vehicles.

4. Federal Court

Even aspects of district court litigation remain opaque to federal judges, if our interviews are representative. Some judges expressed confusion over who represents the agency. The SSA lacks independent litigating authority, so the U.S. Attorney is nominally the government's lawyer.¹⁶⁷ With some exceptions, however, lawyers from the SSA's Office of General Counsel ("OGC") do the heavy lifting on the substance of the appeal.¹⁶⁸ In fact, OGC lawyers seem to prefer to litigate as "special assistant U.S. attorneys," or "SAUSAs,"¹⁶⁹ a designation that enables OGC lawyers to litigate with very little involvement of the U.S. Attorney.¹⁷⁰

OGC lawyers are organized into ten regions, with each region covering litigation in all districts within no more than two circuits.¹⁷¹ In some instances, OGC lawyers within a region specialize in one or two districts.¹⁷² These lawyers gain a deep understanding of that district's rules and standing orders and even of particular judges and their tendencies. One OGC lawyer we spoke with litigates exclusively in front of three judges, enabling her to speak at length about each one's idiosyncrasies.¹⁷³ In other regions, OGC lawyers litigate in different districts, for at

¹⁶⁶ Cf. JERRY L. MASHAW ET AL., ADMINISTRATIVE LAW: THE AMERICAN PUBLIC LAW SYSTEM 541 (7th ed. 2014).

¹⁶⁷ OGC Lawyer 1 at 1; 28 U.S.C. §§ 516-519. See generally Neal Devins & Michael Herz, *The Battle That Never Was: Congress, the White House, and Agency Litigation Authority*, 61 LAW & CONTEMP. PROBS. 205 (1998).

¹⁶⁸ OGC Lawyer 1 at 1; OGC Lawyer 2 at 1; OGC Lawyer 3 at 3; DOJ Official 2 at 1; DOJ Official 3 at 1.

¹⁶⁹ E.g., OGC Lawyer 4 at 1; OGC Lawyer 5, Second Interview at 3; OGC Lawyer 6 at 2; OGC Lawyer 7 at 1; OGC Lawyer 1 at 1. One OGC lawyer we spoke suggested that the degree of AUSA involvement depends on its trust of OGC. OGC Lawyer 8 at 5.

¹⁷⁰ OGC Lawyer 9 at 1.

¹⁷¹ Notice Announcing Addresses for Service of Process, 79 Fed. Reg. 4519-04 (Jan. 28, 2014). Before January 2014, OGC's organization was more haphazard and depended to a greater degree on workload distribution. Lawyers in the Denver Regional Office, for example, handled cases within three different circuits, including the Ninth. Because OGC lawyers in the San Francisco and Seattle regions also handled cases within the Ninth Circuit, three offices had to coordinate when contemplating a Ninth Circuit appeal. OGC Lawyer 2 at 1. Attorneys at OGC's headquarters in Woodlawn, Maryland also handle a share of the agency's program litigation workload.

¹⁷² E.g., OGC Lawyer 10 at 1.

¹⁷³ OGC Lawyer 10 at 4.

least two reasons. OGC lawyers should get the “joy of a good district and the pain of a bad district equally,” we were told.¹⁷⁴ Also, workflow difficulties – appeals can spike in one district and abruptly subside in another – can make specialization by district hard to manage administratively.

Caseload pressures affect how OGC lawyers handle cases. With eleven briefing deadlines looming when we spoke, one OGC lawyer aptly described his workload as “crushing.”¹⁷⁵ Among the OGC lawyers we interviewed, workload estimates ranged from four briefs to a dozen briefs due per month,¹⁷⁶ added to a slate of other duties these lawyers handle for the agency.¹⁷⁷ An OGC lawyer described three days as an “enormous” amount of time to spend writing a single brief.¹⁷⁸

This workload keeps many OGC lawyers from scrutinizing a case in any earnest until several days before a brief is due.¹⁷⁹ This timing in turn has unfortunate implications for requests for voluntary remands, or “RVRs.” Several officials we spoke with described the RVR process as something akin to an Appeals Council backstop. If a case that cannot be defended survives Appeals Council review, OGC lawyers will catch these errors by requesting voluntary remand.¹⁸⁰ The agency typically requests voluntary remand in about 15% of appeals annually. OGC lawyers will rarely seek an RVR in a case based on district court case law, even that authored by the judge before a current appeal is pending.¹⁸¹ As several OGC lawyers told us, the

¹⁷⁴ OGC Lawyer 1 at 6.

¹⁷⁵ OGC Lawyer 11 at 4.

¹⁷⁶ OGC Lawyer 2 at 4 (8-9 briefs per month); OGC Lawyer 3 at 1 (8-10 briefs per month); OGC Lawyer 1 at 2 (4-8 briefs per month); OGC Lawyer 12 at 2 (2-3 briefs per week); OGC Lawyer 13 at 2 (four briefs per month); OGC Lawyer 10 (4-5 briefs per month); OGC Lawyer 9 (five briefs per month); OGC Lawyer 6 (6-8 briefs per month).

¹⁷⁷ These duties include representing the agency in employment litigation, dealing with labor/management issues, and providing advice on other agency programs. *E.g.*, OGC Lawyer 4 at 1.

¹⁷⁸ OGC Lawyer 10 at 1.

¹⁷⁹ *E.g.*, OGC Lawyer 10 at 1; OGC Lawyer 11 at 3; OGC Lawyer 2 at 2; OGC Lawyer 9 at 2.

¹⁸⁰ Agency Official 5 at 3; OGC Lawyer 5, Second Interview at 3.

¹⁸¹ *But see* OGC Lawyer 2 at 2 (indicating that, while district court preferences aren’t supposed to affect RVR thresholds, lawyers do take them into account in extreme circumstances).

Appeals Council, which has to agree to an RVR,¹⁸² will generally not do so if the basis is a circuit decision that the lawyer cannot translate into existing agency policy.¹⁸³

Several of our interview subjects believe that the last-minute review OGC lawyers give cases can discourage RVRs, for at least two reasons. First, RVRs may take substantial work. The OGC lawyer has to write a memorandum for the Appeals Council. While the Appeals Council acquiesces in nearly 95% of RVR requests,¹⁸⁴ the process is cumbersome.¹⁸⁵ Appeals Council personnel also find fault with the process, complaining that briefing deadlines impose difficult time constraints.¹⁸⁶ Rather than jump through these hoops, OGC lawyers may simply file a merits brief and let the court make the call on the remand.¹⁸⁷ Second, we were told, the RVR process can all but require a last minute request for an extension. Appeals Council review can take a week to complete, a significant delay if the OGC lawyer doesn't identify a fatal flaw in an ALJ's order until three days before a brief is due.¹⁸⁸ Rather than frustrate a federal judge with a last minute request for an extension, an OGC lawyer might file the merits brief.¹⁸⁹

Caseload pressures also exacerbate the inconvenience and inefficiencies caused by the bewildering array of local rules and standing orders OGC lawyers and claimant representatives have to navigate. We address the problem of balkanized procedural governance in Part V. For

¹⁸² Agency Official 6, First Interview at 2.

¹⁸³ OGC Lawyer 18 at 4 (observing that the Appeals Council will not grant an RVR based on *Mascio v. Colvin*, 780 F.3d 632 (4th Cir. 2015)); OGC Lawyer 12 at 4 (insisting that it is “about impossible” to get an RVR accepted by the Appeals Council based on *Bird v. Commissioner of Social Security*, 669 F.3d 337 (4th Cir. 2012)); OGC Lawyer 24 at 3 (observing that the Appeals Council will not grant an RVR request if it is based on a “court position”). The situation is different, of course, if the agency has issued an acquiescence ruling for the circuit decision.

¹⁸⁴ Agency Official 6, First Interview at 2.

¹⁸⁵ OGC Lawyer 12 at 2; OGC Lawyer 1 at 6. *Cf.* OGC Lawyer 13 at 1 (describing the RVR process as “administratively complicated”).

¹⁸⁶ Agency Official 6, First Interview at 3.

¹⁸⁷ OGC Lawyer 10 at 2; OGC Lawyer 13 at 1.

¹⁸⁸ One OGC lawyer told us that he does a defensibility analysis throughout his work on a case, including right after the complaint is filed. OGC Lawyer 14 at 1.

¹⁸⁹ One OGC lawyer told us of expressions of disapproval he has received from district judges when he has requested extensions in order to pursue the RVR process. OGC Lawyer 2 at 2.

present purposes, the following examples illustrate the lack of agreement among the districts on procedural rules for social security cases:

- In the Western District of New York, the plaintiff-claimant files the opening brief and the SSA responds. Until June 2015, the order in the Eastern District of New York was the reverse.¹⁹⁰ Both sides move for summary judgment in the District of Minnesota.¹⁹¹ Motions for summary judgment “are not appropriate” in social security cases in the neighboring Northern District of Iowa.¹⁹²
- In the Eastern District of Missouri, the plaintiff must file his brief in support of the complaint thirty days after the Commissioner’s answer is filed. The Commissioner must respond thirty days later, and the reply is due fourteen days after that. Neither of the first two briefs can exceed fifteen pages.¹⁹³ In the Western District of Missouri, the corresponding briefing deadlines are forty days, forty days, and twenty-one days, and the briefs cannot exceed fifteen pages in length, exclusive of a “facts” section.¹⁹⁴
- The Commissioner can file the certified administrative record in lieu of an answer in the Northern District of Illinois.¹⁹⁵ The Southern District of Illinois does not permit this efficiency.¹⁹⁶

These variations reflect the singular nature of disability claims, insofar as they require trial courts to act as appellate tribunals. The square peg of social security appeals has never fit neatly into the round hole of the Federal Rules of Civil Procedure.¹⁹⁷

Governing procedures do not just differ by district boundary. In the Central District of California, the highest volume district for disability appeals in the country,¹⁹⁸ twenty or so discrete case management orders specify procedural requirements that vary by judge. The

¹⁹⁰ *Compare* Standing Order, In the Matter of Actions Seeking Review of the Commissioner of Social Security’s Final Decisions Denying Social Security Benefits, W.D.N.Y., Sept. 6, 2013, *with* Administrative Order 2008-5, In re: Scheduling in Social Security Cases, E.D.N.Y., Mar. 20, 2008. In June 2015, the Eastern District reversed the order to match what the parties do in the Western District. Administrative Order 2015-05, In re: Scheduling in Social Security Cases, E.D.N.Y., June 19, 2015.

¹⁹¹ D. Minn. Local R. 7.2(d).

¹⁹² N.D. Iowa Local R. 56(i).

¹⁹³ E.D. Mo. Local R. 56-9.02.

¹⁹⁴ W.D. Mo. Local R. 9.1.

¹⁹⁵ N.D. Ill. Local R. 8.1(b).

¹⁹⁶ S.D. Ill. Local R. 9.1(b).

¹⁹⁷ Morton Denlow, *Substantial Evidence Review in Social Security Cases as an Issue of Fact*, 2 FED. CTS. L. REV. 99, 127-131 (2007) (reporting results of a survey indicating different procedural approaches to social security appeals).

¹⁹⁸ Communication with OGC Lawyer 15, Sept. 4, 2015.

differences are many and substantial. Some judges require the parties to file a single “joint stipulation” in lieu of merits briefs,¹⁹⁹ for example, while others prefer standard briefing.²⁰⁰ Some judges require the parties to exchange letters discussing settlement before filing a merits brief or stipulation.²⁰¹ Others require an in-person or telephonic meet-and-confer after the plaintiff drafts her merits argument.²⁰² Still others dispense with any mandatory settlement discussions altogether.²⁰³

When a federal judge remands a case, the order goes back to the Appeals Council, where a specialized division receives it and issues the actual remand order to the ALJ.²⁰⁴ These orders are typically boilerplate and incorporate the district court’s decision by reference.²⁰⁵ The Appeals Council usually processes these remands fairly expeditiously,²⁰⁶ although a claimant representative complained to us that his have languished for half a year.²⁰⁷ A year can pass between the district court’s order and a second decision from an ALJ.²⁰⁸ The norm has been to have remands return to the same ALJ who rendered the decision in the first instance.²⁰⁹ Remands typically do not return to the same decision writer.²¹⁰

¹⁹⁹ Case Management Order, *Santibanez v. Astrue*, Civ. No. 11-396, C.D. Cal., Mar. 11, 2011, at 4.

²⁰⁰ Case Management Order, *Darling v. Astrue*, Civ. No. 11-1247, C.D. Cal., Feb. 15, 2011, at 3.

²⁰¹ Case Management Order, *Santibanez v. Astrue*, Civ. No. 11-396, C.D. Cal., Mar. 11, 2011, at 3.

²⁰² Case Management Order, *Baylioni v. Astrue*, Civ. No. 11-1819, C.D. Cal., Mar. 11, 2011, at 6-7.

²⁰³ *See generally* Case Management Order, *Darling v. Astrue*, Civ. No. 11-1247, C.D. Cal., Feb. 15, 2011.

²⁰⁴ Agency Official 6, First Interview at 2.

²⁰⁵ Agency Official 6, First Interview at 2.

²⁰⁶ The agency reported to us that the average age for court remands pending at the end of FY 2015 was 52 days.

²⁰⁷ Claimant Representative 2 at 6.

²⁰⁸ U.S. GOV’T ACCOUNTABILITY OFFICE, 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, Apr. 2007, at 9.

²⁰⁹ *But see* https://aalj.org/system/files/documents/aaj_legal_analysis_3-15-16.pdf (indicating the possibility that the agency will not assign court remands to the same ALJs); *Leading the Hearing and Appeals Process Into the Future*, *supra* note 74 at 11. Under the system presently in operation, some court remands might not return to the same ALJs who rendered decisions the first time. Sometimes ALJs travel to remote hearing sites, where they will hold hearings for a week or two at a time. If a remand comes to one of these hearing sites, it will go to whatever ALJ is assigned to the site at the time. ALJ 10 at 4. Also, a district court can order that a remand go to a different ALJ if the first ALJ manifested bias. *E.g.*, *Card v. Astrue*, 752 F. Supp. 2d 190, 190-193 (D. Conn. 2010). It is “very hard” if not “impossible” to get a district court to make this finding, one claimant representative told us. Claimant

The agency can always appeal the district court’s decision, but it almost never does so.²¹¹ The courts of appeals might receive somewhere in the neighborhood of 650 social security appeals each year, no more than twenty of which are affirmative appeals by the Commissioner.²¹² In FY 2014, the agency filed exactly one appeal.²¹³ Several reasons might explain this low incidence of appeal, but one institutional fact is surely important: the Solicitor General of the United States must sign off on any appeal the SSA might want to take.²¹⁴ A request for permission to appeal goes from an OGC lawyer to supervisors in her office, then to OGC headquarters, then to a line attorney working on the appellate staff at the Department of Justice’s Civil Division, then to the head of the Civil Division, then to a line attorney at the Solicitor General’s office, then to the Deputy Solicitor General, and finally to the Solicitor General.²¹⁵ Consultation with the U.S. Attorney is also necessary. No district judge we interviewed could ever recall an appeal from a remand or reversal order, and only one claimant representative recalled an agency appeal.²¹⁶ A couple of OGC lawyers suggested to us that their colleagues do not bother trying to get permission to appeal at all, out of a sense of futility.²¹⁷

B. The Numbers

Several statistics reveal quite a lot about the cases that arrive in federal court. A first

Representative 1 at 6. If an ALJ has already heard a case twice, then it goes to a different ALJ upon a second remand. HALLEX, *supra* note 38 § I-2-1-55(d)(5).

²¹⁰ Decision Writer 2 at 2.

²¹¹ OGC Lawyer 16, First Interview at 1. An earlier estimate pegged the number a bit higher. GAO, SSA HAS TAKEN STEPS, *supra* note 208 at 10 (observing that “no more than 20 district court cases have been appealed by the agency to appellate courts each year since 2000”).

²¹² Statistical Tables for the Federal Judiciary – June 2014, Table B-7, at 2, at www.uscourts.gov/statistics-reports/statistical-tables-federal-judiciary-june-2014.

²¹³ Communication with OGC, Nov. 17, 2015. In FY 2013, the agency filed six appeals, and as of November 17, 2015, the agency had filed three appeals during FY 2015. *Id.*

²¹⁴ OGC Lawyer 16, First Interview at 1.

²¹⁵ DOJ Official 1 at 1; OGC Lawyer 16, First Interview at 1. *See generally* BERNARD ROSEN, HOLDING GOVERNMENT BUREAUCRACIES ACCOUNTABLE 127-28 (3d ed. 1998).

²¹⁶ Claimant Representative 3 at 3.

²¹⁷ OGC Lawyer 8 at 5; OGC Lawyer 10 at 5.

insight comes from caseloads and decision patterns. Assuming that FY 2014 statistics are reasonably illustrative, they indicate several interesting phenomena. Only about ten percent of ALJ decisions that could potentially produce federal court appeals actually do so.²¹⁸ We discuss the significance of this number in the next Part.

To a significant extent, caseloads also affect workflows at each level of the process. The initial and reconsideration processes take about seven months to complete.²¹⁹ According to Michael Asimow and Jeffrey Wolfe, the current caseload of 500-700 cases per year is about double what the ALJ system was originally designed to handle.²²⁰ On average, ALJs decide about 46 cases per month, although this number has decreased recently, and they spend about two-and-a-half hours on each case.²²¹ The process takes more than four hundred days from the time of filing until the date of the ALJ's decision.²²² Appeals Council analysts on average prepare two action documents per day, and AOs and AAJs generally take about ten "actions" per day. The average processing time for a claim at the Appeals Council is just under a year.²²³

C. Quality Assurance

The sheer number of claims the agency processes each year is breathtaking. Production

²¹⁸ See *infra* note 283 & accompanying text.

²¹⁹ In FY 2014, the average wait time for a decision on an initial application was 110 days, and the average wait time for a reconsideration determination was 108 days. SOCIAL SECURITY ADMINISTRATION, ANNUAL PERFORMANCE REPORT, 2014-2016, at 83-84 (2015).

²²⁰ Michael Asimow & Jeffrey S. Wolfe, *Thinking Outside the APA Box: A New Social Security Tribunal*, ADMIN. & REG. LAW NEWS, Vol. 38, No. 2, at 3 (2013).

²²¹ KRENT & MORRIS, STATISTICAL APPENDIX, *supra* note 67 at 6 (decisions per month); Complaint, Association of Administrative Law Judges v. Colvin, Civ. No. 13-2925, N.D. Ill., Apr. 18, 2013, at 19 (hours per claim). ALJ productivity has been declining recently. OIG, EFFORTS TO ELIMINATE THE HEARING BACKLOG, *supra* note 69 at 4. Our interview subjects reported a figure around two-and-a-half hours. This figure is consistent with what a recent report commissioned by the Association of American Law Judges calculated. Paullin et al., *supra* note 46 at iii; *id.* at 42.

²²² ANNUAL PERFORMANCE REPORT, *supra* note at 219 at 84 (2015). Another source put the average processing time for a claim at the ALJ level at more than 450 days. OIG, EFFORTS TO ELIMINATE THE HEARING BACKLOG, *supra* note 69 at 1.

²²³ OIG, REQUEST FOR REVIEW WORKLOADS AT THE APPEALS COUNCIL, *supra* note 146 at 5.

goals for ALJs have generated criticism that the agency shortchanges decisional quality to boost output, or number of claims adjudicated.²²⁴ ALJs have come under fire in Congress for ostensibly granting benefits excessively as a strategy to deal with its high volume of claims. Both in response to this criticism and proactively, the agency has undertaken a number of initiatives, designed to ensure a threshold of decisional quality at robust production levels.²²⁵

Two important features of the ALJ corps must be kept in mind when evaluating the quality assurance initiatives that the agency has undertaken. First, the agency has limited say over whom it hires as ALJs.²²⁶ It selects ALJs from a list compiled by the Office of Personnel Management, one created for all agencies that employ ALJs and not just for the SSA.²²⁷ Second, the agency is strictly limited in terms of the performance reviews it can conduct of ALJs, as well as the disciplinary measures it can take against those that are underperforming.²²⁸ It cannot sanction an ALJ for her decisions, for example. For a long time low productivity rarely qualified as good cause to justify an ALJ's removal.²²⁹ Recently the agency's power to remove an ALJ for this reason has grown modestly.²³⁰

²²⁴ See generally U.S. House of Representatives, Committee on Oversight and Government Reform, *Misplaced Priorities: How the Social Security Administration Sacrificed Quality for Quantity in the Disability Determination Process*, Staff Report, 113th Cong., Dec. 18, 2014; Mark J. Warshawsky & Ross A. Marchand, *Disability Claim Denied? Find the Right Judge*, WALL ST. J., Mar. 8, 2015; Damian Paletta, *Disability-Claim Judge Has Trouble Saying 'No'*, WALL ST. J., May 19, 2011.

²²⁵ For summaries of some of these quality assurance initiatives, see generally 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 (2015); Statement of Debra Bice, Chief Administrative Law Judge, Social Security Administration, before the Senate Committee on Homeland Security and Government Affairs, Oct. 7, 2013; Statement of Judge Patricia Jonas, Hearing Before the Committee on Homeland Security and Governmental Affairs, Permanent Subcommittee on Investigations, Sept. 13, 2012.

²²⁶ See generally Kent Barrett, *Resolving the ALJ Quandary*, 33 J. NAT'L ASS'N ADMIN. L. JUDICIARY 644, 652-654 (2013).

²²⁷ Vanessa K. Burrows, Cong. Res. Serv., *Administrative Law Judges: An Overview*, Apr. 13, 2010, at 2.

²²⁸ *Id.* at 7-9.

²²⁹ See generally Jeffrey S. Lubbers, *The Federal Administrative Judiciary: Establishing an Appropriate System of Performance Evaluation for ALJs*, 7 ADMIN. L.J. AM. U. 589, 589-593 (1993/1994).

²³⁰ *Shapiro v. Social Security Administration*, 800 F.3d 1332, 1337 (Fed. Cir. 2015).

The SSA began to implement new quality assurance initiatives toward the end of the 2000s.²³¹ They generally involve decisional tools, feedback on decision-making, and the use of sophisticated analytics to identify and fix recurring decisional errors. The former category includes the electronic claims analysis tool, or “eCAT.” Used by all state-level DDS offices since 2012,²³² eCAT is a web-based program that walks a disability examiner through a disability determination, to make sure all relevant issues get addressed and all agency policies honored.²³³ ACAT, or the tool used by the Appeals Council to deconstruct ALJ decisions since 2008,²³⁴ is the equivalent at the Appeals Council. Very recently the Appeals Council began use of “clustering analysis,” whereby ALJ decisions raising the same or similar issues get allocated in batches to the same analysts. Previously an analyst might encounter issues much less frequently. Now, work can be assigned such that an analyst will encounter an issue repeatedly and thereby develop expertise in it.²³⁵ Finally, the Appeals Council began to use natural language processing in 2015. An algorithm can identify certain errors in decisions, such as findings at Step Three and Four of the sequential process that are inconsistent.²³⁶

These decisional tools also include the eBB, designed to assist ALJ decision-making. The eBB enables ALJs to organize exhibits, take notes on them, prepare questions for hearings, assign weight to particular medical opinions, and make preliminary findings for the steps of the

²³¹ Agency Official 3, Fourth Interview at 1.

²³² Agency Official 3, Fourth Interview at 1.

²³³ OFFICE OF THE INSPECTOR GENERAL, SOCIAL SECURITY ADMINISTRATION, QUICK RESPONSE EVALUATION, THE EFFECTS OF THE ELECTRONIC CLAIMS ANALYSIS TOOL, A-01-11-21193, July 2011, at 1; *see also* ANNUAL PERFORMANCE REPORT, *supra* note 219 at 76.

²³⁴ Agency Official 3, Fourth Interview at 1.

²³⁵ Agency Official 3, Fourth Interview at 5-6.

²³⁶ Agency Official 3, Fourth Interview at 6. An agency report of January 2016 reported that the Appeals Council is currently “testing” the use of natural language processing. *Leading the Hearing and Appeals Process Into the Future*, *supra* note 74 at 10.

sequential evaluation process as they review cases.²³⁷ As noted, the eBB also generates instructions for decision writers. It represents the SSA's effort to map out policy complaint decisions, then to chart a pathway for ALJs to follow to accurate and legally supported outcomes.²³⁸ By reducing the use of decisional heuristics that have caused ALJ error, the eBB, at least in theory, can ensure that ALJ's decisions remain consistent with governing law.

Since the agency only began to urge its use in late 2013,²³⁹ the eBB's popularity may well expand, but at present only 15-20% of ALJs use the tool regularly.²⁴⁰ Opinions on its use vary. Of the few ALJs among our interview subjects who utilize the eBB consistently,²⁴¹ one believes that it has "absolutely made [her] a better judge." The tool makes sure she answers all of the questions she needs to address for a complete decision.²⁴² Another user described the tool as "cumbersome,"²⁴³ however, and the third reported diminished productivity due to his eBB use.²⁴⁴ Among ALJs who do not use the eBB, some criticized it as too time-consuming²⁴⁵ or confusing to use,²⁴⁶ and others doubt that it contributes more than marginally to decisional quality.²⁴⁷ Most decision writers we interviewed gave eBB-generated instructions negative reviews,²⁴⁸ although one praised eBB-generated instructions as more complete than others.²⁴⁹

²³⁷ Misc. Agency Officials at 1-2; *see also* Statement of Debra Bice, Chief Administrative Law Judge, Social Security Administration, Before the Senate Committee on Homeland Security and Government Affairs, Oct. 7, 2013; ANNUAL PERFORMANCE REPORT, *supra* note 219 at 76.

²³⁸ Agency Official 3, Third Interview at 1.

²³⁹ ALJ 22 at 1.

²⁴⁰ Misc. Agency Officials at 1.

²⁴¹ ALJ 21 at 3; ALJ 13 at 1; ALJ 22 at 1. A couple of ALJs use the eBB either occasionally or for some but not all of their decision making. ALJ 6 at 2 (occasionally); ALJ 3 at 4 (takes notes on eBB).

²⁴² ALJ 13 at 1.

²⁴³ ALJ 22 at 1; *see also* ALJ 23 at 2 (using the same term).

²⁴⁴ ALJ 21 at 3.

²⁴⁵ ALJ 7 at 8; ALJ 2 at 1;

²⁴⁶ ALJ 7 at 8 (time-consuming); ALJ 2 at 1 (time-consuming); ALJ 16 at 3 (confusing); ALJ 5 at 4 (confusing).

²⁴⁷ ALJ 19 at 2.

²⁴⁸ One decision writer described her problem with eBB instructions colorfully. They are like a "teenager's bedroom: you know that everything is in there, but how do you find a scarf?" Decision Writer 2 at 3. *See also* Decision Writer 8 at 3; Decision Writer 9 at 2. An ALJ complained that, whereas he prefers to take about two pages of notes for a hearing, the eBB generates 8-12 pages, unnecessarily in his view. ALJ 23 at 2.

²⁴⁹ Decision Writer 7 at 1.

ALJs get feedback from a number of sources. The first are remands themselves, whether from district courts or the Appeals Council. In 2010, the Appeals Council also began performing “pre-effectuation reviews,” or “own motion review” of non-appealed ALJ decisions. The Appeals Council can select a decision to consider reviewing within a short period of time after its issuance.²⁵⁰ The Appeals Council cannot select decisions for such review based on a specific ALJ’s identity or a specific hearing office; it can choose cases based on selective sampling, such as type of claimed impairment.²⁵¹ In these cases, the Appeals Council can choose not to grant review and thereby effectuate the case without changes, or it can grant review and issue a new decision or remand the case to an ALJ for further action.

ALJs can also get feedback through focused reviews. Data captured at ODAR can help the agency identify outlier judges. The SSA can then undertake a “focused review” of closed cases to determine what might be causing problems, if they exist. In a focused review, a special unit of specially trained attorneys at the Appeals Council review a sample of at least sixty closed decisions for error, whether appealed or not. The results of these studies are reported to a Board of Executives, which then decides what individual guidance to provide to the particular ALJ.²⁵²

The SSA selects all sorts of individuals, including decision writers, vocational experts, and physicians, for focused reviews. It can also conduct a focused review on a type of impairment.²⁵³ But of the 20-25 focused reviews it conducts per year, 90% are of ALJs.²⁵⁴ The agency does not necessarily select ALJs for focused reviews because of outlier decision patterns, but of the fifty focused reviews completed by mid-2014, thirty involved ALJs with allowance

²⁵⁰ 20 C.F.R. §§ 404.969 and 416.1469.

²⁵¹ 20 C.F.R. §§ 404.969(b)(1) and 416.1469(b)(1).

²⁵² Agency Official 3, Fourth Interview at 2.

²⁵³ Agency Official 3, Fourth Interview at 2.

²⁵⁴ *Id.*

rates over 75%.²⁵⁵ This emphasis comes as little surprise, given Congressional scrutiny of an ostensibly profligate agency.²⁵⁶ From 2007 to 2013, the overall ALJ allowance rate dropped, while the number of ALJs with extremely low rates remained stable.²⁵⁷ Agency officials insist that this change reflects more policy-compliant decision-making. Several claimant representatives told us that they believe that ALJs have become increasingly stingy with benefits.²⁵⁸

A final feedback initiative involves a tool called “How MI Doing,” a web-based program that provides an ALJ with information about numbers of cases she has decided, the rates at which the Appeals Council remands her decisions (the “agree” rate), and the top several reasons for her remands.²⁵⁹ How MI Doing doesn’t disaggregate remands by the Appeals Council and the district courts, although an ALJ can access a spreadsheet of all of her remands, click on a docket number, and bring up the remand order. The data enable an ALJ to compare herself to other ALJs in the same office and region, as well as nationally.

The Office of Disability Adjudication and Review, or ODAR – the arm of the agency that houses ALJs and the Appeals Council – pursues most of these quality initiatives. Also worth discussing are a couple of endeavors that OGC has pursued to improve the representation the

²⁵⁵ U.S. House of Representatives, Committee on Oversight and Government Reform, *Misplaced Priorities: How the Social Security Administration Sacrificed Quality for Quantity in the Disability Determination Process*, Staff Report, 113th Cong., Dec. 18, 2014, at 13 n.31.

²⁵⁶ See *infra* note 224.

²⁵⁷ Krent & Morris, *supra* note 1 at 380; Statement of Glenn Sklar, Deputy Commissioner, Office of Disability Adjudication and Review, Social Security Administration, Before the House Committee on Oversight and Government Reform, Nov. 9, 2013, at http://www.ssa.gov/legislation/testimony_111913.html (indicating in Figure 1 that the number of high allowance rate judges has dropped but the number of low allowance rate judges has increased from FY 2007 to FY 2013).

²⁵⁸ One claimant representative speculated that public and agency pressure have pushed ALJs in an anti-claimant direction over the past five years. Claimant Representative 4 at 3. Another claimant representative insisted that the national ALJ corps has “definitely . . . shifted to more of an anti-claimant point of view” recently, chalking up the change to training. Claimant Representative 7 at 1. A third claimant representative reported an ALJ telling him “off the record” that ALJs are instructed not to pay certain types of claims. Claimant Representative 8 at 3. A fourth insisted that harsh public scrutiny of ALJs has put them in “duck and cover mode,” and that this pressure has produced more denials. Claimant Representative 9 at 1; *id.* at 2 (explaining the rise of denials by insisting that ALJs don’t want to be the targets of congressional inquiries).

²⁵⁹ Misc. Agency Officials at 2.

agency receives in federal court. The federal judges we spoke with are generally satisfied with the quality of OGC briefs, with a couple of exceptions.²⁶⁰ OGC lawyers in several districts nonetheless have taken it upon themselves to improve their work product.²⁶¹ Among cases filed in 2008 in the District of Maryland, for example,²⁶² data that the agency provided us show that claimants won remand in 70% of appeals.²⁶³ To reduce this rate, an OGC lawyer formed a team of colleagues to focus on District of Maryland litigation. These lawyers met frequently to strategize over issues. If an argument about a particular issue proved persuasive to one judge, the team would then use this argument when the issue next arose. The team leader assembled all District of Maryland decisions and sent memos on them to his colleagues. They created a centralized peer review system for briefs. The OGC lawyers also reached out to Maryland ALJs, to counsel them about recurring issues.²⁶⁴ Among those appeals filed just two years later, in 2010, the District of Maryland remand rate had fallen to 38%, and for cases filed between 2011 and 2014, it has stayed between 34 and 42%.

* * *

“[T]he social security hearing system is probably the largest adjudication agency in the western world,”²⁶⁵ and its complexities are infinite. Two final facts, however, are essential to understanding the disability claims adjudication system as it presently exists. First, a huge number of people file for disability benefits each year, and this population has grown remarkably

²⁶⁰ Most of the federal judges we interviewed expressed this view. One judge, however, criticized the quality of lawyering as “poor.” This judge complained that agency lawyers rely too much on boilerplate in their briefs and do not provide enough specific guidance to record evidence. Federal Judge 13 at 1.

²⁶¹ An OGC lawyer involved with one of these efforts told us that lawyers in his office pursued them on their own, and that the agency more centrally did not organize or direct them. OGC Lawyer 19 at 2.

²⁶² OGC lawyers who litigate in the Northern District of Georgia pursued a similar strategy, to similar apparent effect. OGC Lawyer 17.

²⁶³ See Data Appendix for a description of these data.

²⁶⁴ OGC Lawyer 17 at 3; OGC Lawyer 18 at 8.

²⁶⁵ *Barnhart v. Thomas*, 540 U.S. 20, 28-29 (2003) (internal quotations and alterations omitted).

over the past fifteen years.²⁶⁶ The Great Recession witnessed a 20% spike in the number of claimants and a 48% increase in the number of hearing requests.²⁶⁷ This rise, an increased emphasis on decisional quality, and a decrease in the number of ALJs for part of this time, have meant that one million cases presently await a hearing.²⁶⁸

Second, the agency must grapple with this ever-increasing workload in a hothouse political environment, one that subjects the agency to continuous, intense, and sometimes confounding scrutiny.²⁶⁹ In recent years, for example, members of Congress have chastised the agency for what they perceive as excessive generosity with benefits.²⁷⁰ As the agency responds to this oversight and other influences, its relative priorities may adjust. In 2011, for example, the Commissioner insisted that “[e]liminating our hearings backlog and preventing its recurrence remains our number one priority.”²⁷¹ More recently, the agency has paid more heed to decisional quality. We mention this oversight not to suggest that it is misplaced, but to help federal judges make further sense of what the agency tries to accomplish and why.

²⁶⁶ In 2000, 1.3 million claimants filed for benefits. By 2014, that number had nearly doubled, to 2.5 million. See www.ssa.gov/oact/STATS/dibstat.html.

²⁶⁷ *Id.*; see also Stephen C. Goss, Testimony to the Senate Committee on Finance, *The Foreseen Trend in the Cost of Disability Insurance Benefits*, July 24, 2014, at www.ssa.gov/OACT/testimony/SenateFinance_20140724.pdf; William R. Morton, Cong. Res. Serv., *Social Security Disability Insurance (SSDI) Reform: An Overview of Proposals to Manage the Growth in SSDI Rolls*, Jan. 9, 2015, at 15; OIG, EFFORTS TO ELIMINATE THE HEARING BACKLOG, *supra* note 69 at 3.

²⁶⁸ OIG, EFFORTS TO ELIMINATE THE HEARING BACKLOG, *supra* note 69 at 4; *id.* at 12.

²⁶⁹ *E.g.*, David Fahrenthold, *The Biggest Backlog in the Federal Government*, WASH. POST, Oct. 18, 2014. In the mid-2000s, the agency was criticized for a “culture of denial.” *E.g.*, *Failing the Disabled*, CBS NEWS, Jan. 15, 2008. Several years later, the narrative shifted, to a story of “paying down the backlog.” Stephen Ohlemacher, *Judges Tell Lawmakers They are Urged to Approve Social Security Disability Claims*, WASH. POST, June 27, 2013. The empirical evidence for the “paying down the backlog” claim is thin. *E.g.*, Robert Nakosteen & Michael Zimmer, *Approval of Social Security Disability Appeals: Analysis of Judges’ Decisions*, 46 APPLIED ECONOMICS 2783, 2791 (2014) (concluding that “volume of decisions does not appear to directly affect leniency, after controlling for state variables and judicial experience”).

²⁷⁰ *E.g.*, Letter from Three Members of the U.S. House of Representatives Committee on Oversight and Government Reform to the Hon. Carolyn Colvin, July 1, 2014; Staff Report, U.S. House of Representatives, *Systemic Waste and Abuse at the Social Security Administration: How Rubber-Stamping Disability Judges Cost Hundreds of Billions of Taxpayer Dollars*, 113th Cong., June 10, 2014.

²⁷¹ Statement of Michael J. Astrue, Commissioner, Social Security Administration, Before the House Committee on Ways and Means, Subcommittee on Social Security and the House Committee on the Judiciary Subcommittee on the Courts, Commercial and Administrative Law, July 11, 2011.

Part III. The Federal Court Remand Rate and What It Means

The federal courts ruled for disability claimants in 45% of the 18,193 appeals they decided in FY 2014, either remanding cases to the agency for further proceedings or remanding and ordering the payment of benefits.²⁷² At first blush, this rate might seem quite high.²⁷³ Certainly agency officials believe that it is. Applicable law requires deference to the ALJ's findings, they argue. Also, most claims arrive in district courts after four layers of internal agency review, and quality assurance initiatives have improved agency decision-making. Claimant representatives disagree. If anything, they insist, the federal courts have given the agency too much credit in recent years.²⁷⁴

Who is right? Do claimants prevail too often in the federal courts? Too rarely? These questions resist any conclusive answer, as we argue in this Part. After putting the remand rate in proper perspective, we explain why the remand rate, viewed statically, says little meaningful about the strictness of judicial review or the quality of agency decision-making. Surely claims adjudication, whether in the agency or in the courts, suffers from problems that need correcting. The SSA and the federal courts can take important steps that could improve agency decision-making and better incentivize federal judges to respect agency goals and priorities. But, to a significant extent, remands result from conflicts between two institutions, the federal courts and the agency, whose design, commitments, and priorities do not fit easily together. Only Congress

²⁷² These numbers are reported on page 143 of the agency's FY 2016 Congressional Justification. *See* www.ssa.gov/budget/FY16Files/2016FCJ.pdf (last visited Dec. 14, 2015). This official remand rate includes voluntary remands. In contested cases, the district courts actually remand a somewhat lower rate. If the RVR rate is 15% and the overall rate is 45%, then the contested rate is only $(45-15)/(100-15)=0.35$, or 35%.

²⁷³ ROBERT J. HUME, *HOW COURTS IMPACT FEDERAL ADMINISTRATIVE BEHAVIOR* 19 (2009) (reporting success rates for agencies in the federal courts); David Zaring, *Reasonable Agencies*, 96 VA. L. REV. 135, 171 (2010) (reporting that federal courts affirm between 58-76% of agency decisions they review); Bert I. Huang, *Lightened Scrutiny*, 124 HARV. L. REV. 1109, 1134-35 (2011) (reporting an overall appellant win rate of about 10-15% in the various federal circuits).

²⁷⁴ *E.g.*, Claimant Representative 6 at 2; Claimant Representative 9 at 2.

can ultimately resolve this fundamental and otherwise indelible institutional tension that, if left alone, seems likely to ensure that claimants win remands in droves.

A. The Remand Rate in Perspective

The FY 2014 remand rate of 45% needs to be put into perspective. History offers one useful metric for comparison. Within six years of the disability benefits program's start, remand rates had climbed to 35%.²⁷⁵ From 1970-1975, federal courts remanded or reversed 45% of appeals,²⁷⁶ although the rate fell steeply by the end of the 1970s.²⁷⁷ The agency's aggressive implementation of continuing disability review, the process by which beneficiaries are reviewed to determine if they remain disabled, met strong headwinds in the federal courts,²⁷⁸ and the remand rate soared to 60.5% by 1984.²⁷⁹ After a decline in the late 1980s, the rate again rose sharply, and by 2001, federal judges reversed or remanded 68% of

²⁷⁵ Rowland, *supra* note 23 at 42 (documenting a reversal rate of 35.3%).

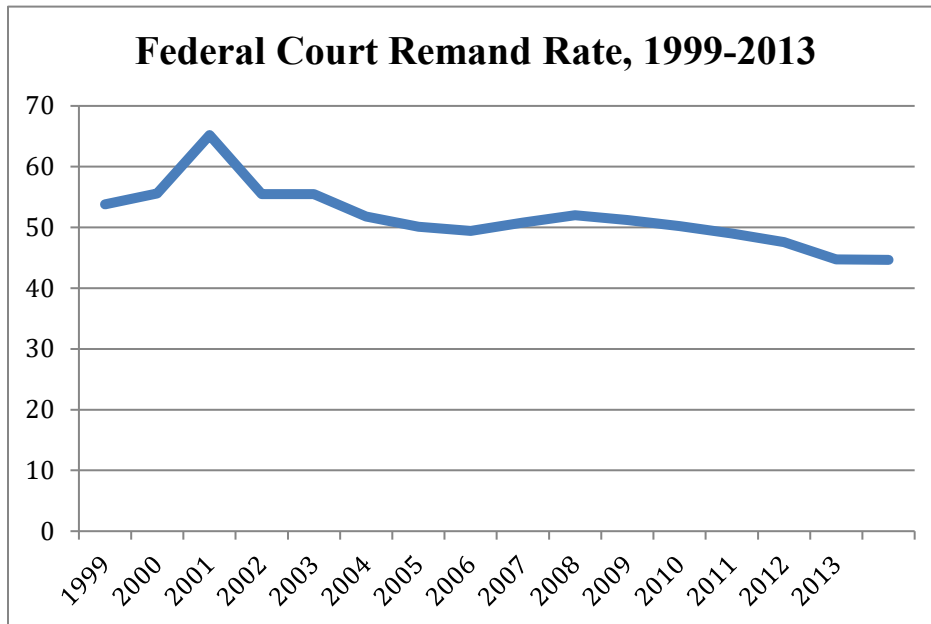
²⁷⁶ MASHAW ET AL., *supra* note 28 at 125.

²⁷⁷ DERTHICK, AGENCY UNDER STRESS, *supra* note 35 at 145.

²⁷⁸ *Id.* at 145-146; Leonard S. Rubenstein et al., *Protecting the Entitlements of the Mentally Disabled: The SSDI/SSI Battles of the 1980s*, 11 INT'L J. LAW & PSYCH. 269 (1988). *E.g.*, Lopez v. Heckler, 713 F.2d 1432, 1441 (9th Cir. 1983) (Pregerson, J., concurring) (denouncing the agency as refusing to honor "the rule of law, the doctrine of separation of powers imbedded in the constitution, and the tenet of judicial supremacy"); Hillhouse v. Harris, 547 F. Supp. 88, 91 (W.D. Ark. 1982); Gerald W. Heaney, *Why the High Rate of Reversals in Social Security Cases?*, 7 HAMLINE L. REV. 1, 11-16 (1984); Jack B. Weinstein, *Equality and the Law: Social Security Disability Cases in the Federal Courts*, 35 SYRACUSE L. REV. 897, 917 (1984).

²⁷⁹ DERTHICK, AGENCY UNDER STRESS, *supra* note 35 at 145.

cases appealed to them. This figure fell to 54% by 2005 and continued to drop.²⁸⁰



The current remand rate, then, hardly stands out as exceptional. Indeed, the agency now fares better in the federal courts than it has since the late 1990s. But the last twenty years have lacked the sorts of steep declines in the remand rate that the agency enjoyed at the end of the 1970s and the 1980s. Struggles over continuing disability review in the 1980s cost the agency significant credibility with the federal courts.²⁸¹ Perhaps a culture of distrust took root then, one that has proven difficult to eradicate.²⁸² If this distrust flows more from historical memory than present concerns about decisional quality, it may be pushing the remand rate to an unwarranted level.

Viewed from a different vantage point, the remand rate seems modest. Comparing reversal and remand rates at various stages in the claims adjudication process is challenging, as delays make unclear which decisions are under review by which decision makers at any

²⁸⁰ These numbers come from the agency's Office of General Counsel and are on file with the authors.

²⁸¹ DERTHICK, *AGENCY UNDER STRESS*, *supra* note 35 at 145.

²⁸² One federal judge with whom we spoke suggested that this is the case, as did an OGC lawyer. Federal Judge 14 at 1; OGC Lawyer 20 at 3.

particular time. Using FY 2014 numbers as illustrative, ALJs can expect claimants to file federal court appeals in under 10% of the decisions adverse to claimants that they issue, and ALJs endure federal court remands for only about 3% of these decisions.²⁸³ About a half-a-dozen court remands each year return to an ALJ on average, in other words.²⁸⁴

The fact that federal judges vacate only 3% of potentially reversible ALJ decisions hardly suggests an excessive remand rate. In fact, this small per capita figure might indicate the problem of too *few* remands. The lengthy path a claim has to follow from a hearing office to a district court, a journey that can take more than two years, might discourage all but the most dogged of claimants.²⁸⁵ Perhaps claimants worn down by the multiple layers of review simply give up on meritorious claims.²⁸⁶

Another perspective on the overall remand rate involves their cost. If a claimant seeking disability insurance benefits ultimately prevails, she can expect to receive about \$1,150 in cash benefits per month.²⁸⁷ This amount barely surpasses the federal poverty line. On average,

²⁸³ Congressional Justification, *supra* note 272 at 143. The agency provided us with the 10% figure. The 3% figure is a rough estimate, derived from FY 2014 numbers. ALJs deny and dismiss claims in about 290,000 cases each year. Of these, the Appeals Council reverses in 18,500 and leaves 32,000 undecided annually. There is a pool of about 239,500 ALJ decisions that federal courts could potentially reverse. Of these, federal courts rule for claimants in 8,200 of cases. Another calculation indicates the small overall impact court remands have on ALJ workloads. In FY 2014, the 1270 ALJs available for decision-making decided 530,574 cases. *See* OIG, EFFORTS TO ELIMINATE THE HEARING BACKLOG, *supra* note 69 at 5 (number of available ALJs in FY 2014); Congressional Justification, *supra* note 272 at 143 (number of cases). 8,200 court remands is 1.5% of this overall workload.

²⁸⁴ If the federal courts rule for claimants in 8,200 cases, and if the agency's ALJ corps includes about 1300 judges, then a little more than six decisions per ALJ are remanded or reversed.

²⁸⁵ Several federal judges expressed this concern, and claimant representatives insisted to us that some of their clients stop short of federal court for reasons unrelated to the merits of their claims. A claimant representative explained that he might advise his client simply to file a new claim rather than appeal to federal court if he believes that the ALJ to whom the case would return is inalterably opposed to his client. Claimant Representative 11 at 2. Another noted that Social Security Ruling 11-1p requires a claimant to choose between filing a new claim for benefits and pursuing an existing claim on appeal while the claim is pending within the agency. Claimant Representative 12 at 2; *see also* Soc. Sec. Ruling 11-1p, 76 Fed. Reg. 45309 (July 28, 2011).

²⁸⁶ As we discuss further in Part IV, such deliberate selection by claimants raises substantial hurdles for any analysis seeking to uncover the causal determinants of the remand rate.

²⁸⁷ This figure refers to a claimant seeking SSDI benefits, not SSI benefits. Testimony of Carolyn W. Colvin, Acting Commissioner, Social Security Administration, Regarding Oversight of Federal Disability Programs, Before the Oversight and Government Reform Committee, U.S. House of Representatives, June 11, 2014. This figure does not include the value of Medicare coverage that a beneficiary would also receive.

however, each grant of benefits costs the federal treasury roughly \$270,000 over the lifetime of a claimant,²⁸⁸ and disability payments in the aggregate at present exceed the amount the federal government pays to service the national debt.²⁸⁹ In 2007, the Government Accountability Office determined that ALJs eventually grant benefits to 66% of those claimants whose claims were remanded from district courts.²⁹⁰ District court remands therefore add a good deal to the national disability benefits tab. But this fact alone hardly indicts federal court review, absent some reason to believe that district court remands produce more erroneous grants of benefits (false positives) than they correct erroneous denials of benefits (false negatives).

Costs also include ALJ resources. Remands take more time for ALJs to adjudicate than claims in the first instance. Assuming they require double the effort, federal court remands add the equivalent of a week or so of work to the average ALJ's docket.²⁹¹ This number strikes us as modest, although we recognize that the enormous and growing docket of claims awaiting ALJ adjudication makes any extra work significant.²⁹² Again, though, unless these cases generate an imbalance of false positives, the added delay that they impose must be weighed against the benefit to thousands of claimants and to the United States of the more accurate implementation of federal disability policy. A week of delay hardly seems a steep price to pay for thousands of corrected benefits decisions.

²⁸⁸ David H. Autor, *The Unsustainable Rise of the Disability Rolls in the United States: Causes, Consequences, and Policy Options*, Nov. 23, 2011, at 4, at <http://economics.mit.edu/files/6880>. Two commentators estimated that the wrongful allowance of benefits from 2005-2014 will cost the federal treasury \$72 billion. Mark J. Warshawsky & Ross A. Marchand, *Reforming the System of Review by Administrative Law Judges in Disability Insurance*, Mercatus Working Paper, Sept. 2015, at 3.

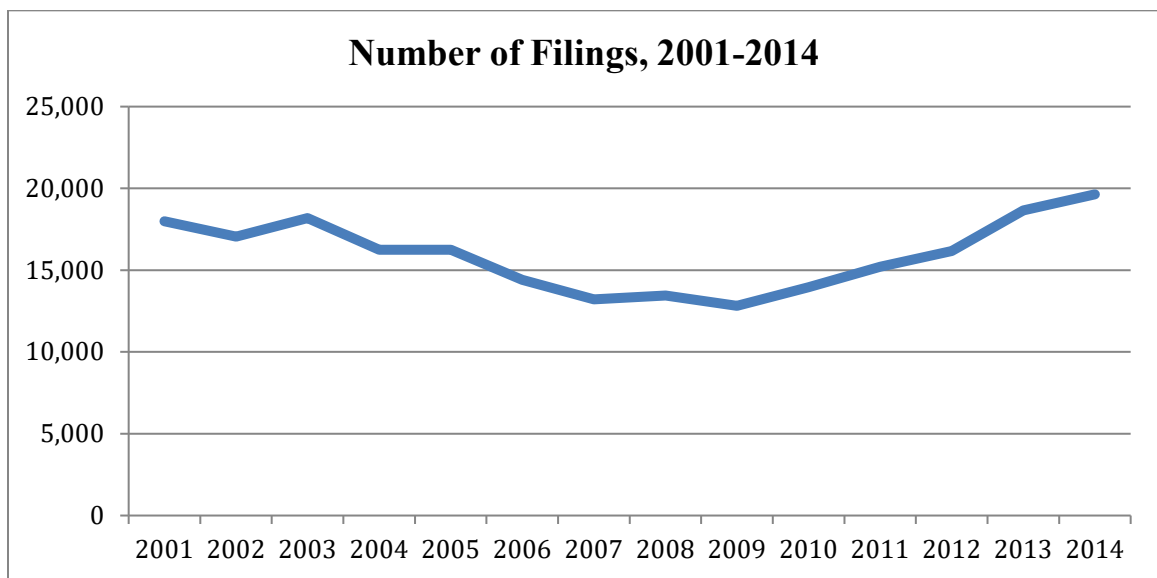
²⁸⁹ Autor, *supra* note 288 at 3.

²⁹⁰ GAO, SSA HAS TAKEN STEPS, *supra* note 208 at 16.

²⁹¹ This is our estimate, based on our calculation that the average ALJ will have six cases remanded by courts per year. (Of course, remands are likely not evenly distributed.) Assuming that the complexity of these cases means that each requires twice the time of a case heard in the first instance, court remands add the equivalent of a dozen cases to each ALJ's docket. The average ALJ decides about 45-48 cases per month. KRENT & MORRIS, STATISTICAL APPENDIX, *supra* note 67 at 7.

²⁹² As of March 2015, the agency had one million claims awaiting ALJ decisions. OIG, EFFORTS TO ELIMINATE THE HEARING BACKLOG, *supra* note 69 at 1.

From a final perspective, however, the 45% remand rate is unusual, if not necessarily troubling. Social security appeals to district courts are *sui generis*, so there isn't a good comparative benchmark, but administrative agency appeals in the federal circuits are a useful if imperfect one.²⁹³ In general, the circuits affirm more than 80% of the administrative agency decisions appealed to them, so that over all plaintiffs win about one in five of their administrative appeals to the circuits.²⁹⁴ Social security plaintiffs therefore prevail at more than double this rate.²⁹⁵ Moreover, while the numbers of appeals might seem insignificant relative to the agency's overall caseload, they are quite large from that of the federal courts, and they are growing:²⁹⁶



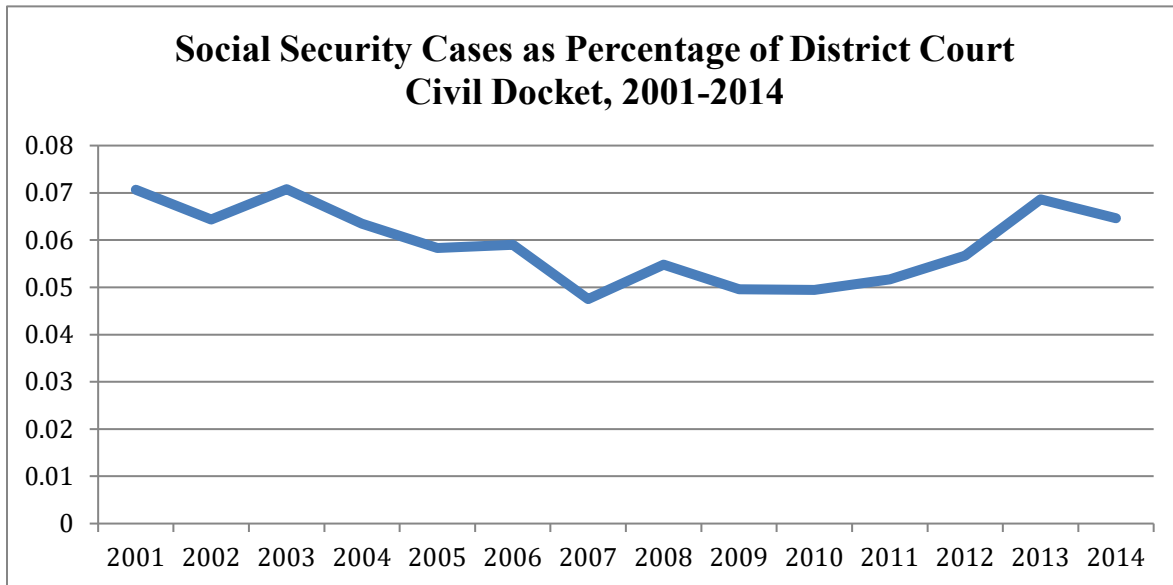
²⁹³ These are imperfect because the range of issues the circuits can review and the applicable standard of review may differ.

²⁹⁴ Statistical Tables for the Federal Judiciary – June 2014, Table B-5, at 1, *at* www.uscourts.gov/statistics-reports/statistical-tables-federal-judiciary-june-2014. In general, appellants win 10-15% of the time in the circuits. Huang, *supra* note 273 at 1134-1135.

²⁹⁵ We emphasize again that the problem of selection in case composition is substantial. Accordingly, we do not think it appropriate to regard the comparison in the text just above as in any way providing evidence as to the causal determinants of the difference in plaintiffs' win rates in administrative appeals. However, these statistics do suggest that either case composition or adjudication, or both, differ substantially in the arenas compared above.

²⁹⁶ *See, e.g.*, U.S. District Courts – Civil Cases Commenced, by Nature of Suit and District, During the 12-Month Period Ending March 31, 2014, Table C-3, *at* www.uscourts.gov/statistics/table/c-3/federal-judicial-caseload-statistics/2014/03/31. The numbers in the chart reflect cases filed during the twelve month period ending March 31 of each year. All data are available at www.uscourts.gov/statistics-reports/analysis-reports/federal-judicial-caseload-statistics.

Perhaps the increase in filings suggests a magnet effect, whereby plaintiff-friendly federal courts attract too many cases. On the other hand, social security cases as a percentage of district courts' civil dockets, while rising as of late, has not spiked:



On the surface, at least, the federal courts seem to be no more of a magnet for appeals than they have been over the past fifteen years.

B. The Strictness of District Court Review

If district courts review agency decisions too strictly, then it seems likely that the benefits of review do not outweigh the costs related to delay and caseload pressures that these appeals create. Agency officials insisted to us that this is so. Even if a 45% remand rate does not stand out once placed in historical or institutional context, they maintained, it is inconsistent with a standard of review that requires deference to agency findings of fact.²⁹⁷ Moreover, the numbers do not seem to reflect judicial awareness for the fact that a claim passes through multiple levels of review within the agency, ones that have benefited recently from various quality assurance

²⁹⁷ For a version of this argument, see Paul R. Verkuil, *An Outcomes Analysis of Scope of Review Standards*, 44 WM. & MARY L. REV. 679, 689 (2002); *id.* at 705-09.

initiatives. At this moment, neither argument persuades us, although as time passes quality assurance initiatives should improve the average quality of potential cases for district court review.

1. The Quality of Decisions Appealed

In a 2007 memorandum, the agency's Chief Administrative Law Judge blamed "most" of the agency's losses in the federal courts on ALJ decisions that "did not comply with our own policy."²⁹⁸ Since then, the agency has hired hundreds of new ALJs, and it has pursued a number of quality assurance initiatives to improve decision-making. Agency officials believe that these efforts have improved claims adjudication, and that the federal courts have failed to appreciate this change.²⁹⁹

Initiatives like the eCAT, the eBB, more targeted training, and focused reviews have succeeded, agency officials believe, citing several indicators.³⁰⁰ Appeals Council remands have dropped from 21.77% in FY 2010 to 14.34% in FY 2014,³⁰¹ even as ALJ allowance rates have also fallen from near 70% in FY 2010 to about 45% now.³⁰² The variance among ALJ grant rates has decreased, compressing differences in decision-making among ALJs.³⁰³ If decisional

²⁹⁸ Letter from Frank A. Cristaudo, Chief Administrative Law Judge, to Colleagues, Dec. 19, 2007, at 3 (on file with authors) ("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. I would like the federal judges to defer to our decisions because we are the experts in Social Security Law; however, they cannot do so if our decisions do not comply with Agency policy.").

²⁹⁹ Agency Official 3, Third Interview at 4. The remand rate had trended steadily downward from FY 2010 to FY 2013, but then it rose in FY 2014. *See* www.socialsecurity.gov/appeals/DataSets/AC05_Court_Remands_NCC_Filed.html.

³⁰⁰ *E.g.*, Ray & Lubbers, *supra* note 225 at 1604-1607.

³⁰¹ *See* www.socialsecurity.gov/appeals/DataSets/AC03_AC_Remands_All_Dispositions.html.

³⁰² OFFICE OF THE INSPECTOR GENERAL, SOCIAL SECURITY ADMINISTRATION, AUDIT REPORT, SUBSEQUENT APPELLATE ACTIONS ON DENIALS ISSUED BY LOW-ALLOWANCE ADMINISTRATIVE LAW JUDGES, A-12-13-13084, July 2014, at 3; *see also* www.disabilityjudges.com/state.

³⁰³ Statement of Debra Bice, Chief Administrative Law Judge, Social Security Administration, before the Senate Committee on Homeland Security and Government Affairs, Oct. 7, 2013.

quality is improving while the remand rate remains stagnant, so the logic goes, then the standard of review in federal court must be increasingly strict.³⁰⁴

This reasoning, at least at the current time, is unconvincing. First, the composition of disability claims likely changed since the Great Recession's onset in 2008. The number of initial applications spiked in 2009, perhaps driven upward by unemployed workers pursuing benefits in lieu of wages.³⁰⁵ This increase continued through 2012.³⁰⁶ Claims motivated by unemployment and not incapacitating impairment are necessarily weaker, if not spurious.³⁰⁷ One would expect lower ALJ allowance rates as the quality of applications declines. If ALJ decisions do not fully correct for this compositional effect, or if some of these rejected claimants with weaker claims appeal, then Appeals Council remand rates also would likely fall.³⁰⁸

³⁰⁴ A number of claimant representatives have noticed changes in ALJ decisions, but very few believe that agency decision-making has improved. One claimant representative acknowledged that ALJ decisions have gotten "more tight" over the past 6-7 years, but that ALJs are now denying claims that shouldn't be denied. Claimant Representative 10 at 3; *see also* Claimant Representative 8 at 3 (acknowledging that ALJ decisions have improved with the FIT, but doubting that agency decision-making has improved overall). Several claimant representatives insisted that agency decision-making has declined in quality. Claimant Representative 7 at 2; Claimant Representative 6 at 2; Claimant Representative 2 at 2. A couple of the federal judges with whom we spoke have noticed modest improvements for the better. A law clerk working for one federal judge claimed that she has noticed the agency's effort "to set forth a more cogent, reviewable document." Law Clerk 1 at 3. Another clerk made a similar observation. Law Clerk 2 at 1. One federal judge told us that he sees good decisions more often now, but that they tend to come from the same ALJ over and over again. Federal Judge 10 at 3. Another federal judge told us that the agency has done a better job "cover[ing] up its tracks," but that "substantively the work is not there." Federal Judge 15 at 2. A number of federal judges have not noticed any changes. *E.g.*, Federal Judge 13 at 1; Federal Judge 12 at 1; Federal Judge 3 at 4; Federal Judge 6 at 1; Federal Judge 16 at 2; Federal Judge 18 at 1; Federal Judge 17 at 1.

³⁰⁵ *See* www.ssa.gov/oact/STATS/dibStat.html.

³⁰⁶ ANNUAL PERFORMANCE REPORT, *supra* note 219 at 81.

³⁰⁷ *E.g.*, Kalman Rupp, *Factors Affecting Initial Disability Allowance Rates for the Disability Insurance and Supplemental Security Income Programs: The Role of Demographic and Diagnostic Composition of Applicants and Local Labor Market Conditions*, SOC. SEC. BULLETIN, Vol. 72, No. 4, at 11 (2012).

³⁰⁸ To understand this point, suppose that there are N benefit claims filed among claimants within the jurisdiction of a given hearing office in a given year. Of these, N_H are high quality claims that should be allowed, while N_0 are meritless claims that would be denied if DDS and/or ALJ adjudication were perfectly accurate. Assume that the fractions D_H and D_0 of these types of claims are actually denied by ALJs. Let the appeal rate following ALJ denial be a_H and a_0 . Then the number of high-quality cases appealed to the Appeals Council level will be $A_H = a_H D_H N_H$, while the number of meritless cases appealed will be $A_0 = a_0 D_0 N_0$. Suppose that the business cycle in a given year affects neither the number of high quality claims filed that year, N_H ; the denial rates D_H and D_0 (these rates capture the behavior of ALJs, rather than applicants); nor the appeal rates a_H and a_0 . Now consider a year in which an economic downturn leads to an increase in N_0 , the number of meritless claims. If we were to follow the cohort of claims filed that year through the agency's claims adjudication process, we would find a typical number of high quality appealed cases, A_H , since (by assumption) none of the variables that determine A_H are affected by the

Second, the decrease in the Appeals Council remand rate does not necessarily indicate across-the-board improvements in ALJ decision-making, a claimant representative insisted to us. He cited a 2014 report by the agency's Office of the Inspector General that indicates that the Appeals Council did not reverse or remand low-allowance ALJs significantly more than ALJs generally.³⁰⁹ This report relied on FY 2010 data; more recent data might reveal a different pattern. But this concern about a time lag goes to a third issue. The agency's quality assurance initiatives have taken time to implement. The current pool of ALJ decisions before the district courts probably does not include those that reflect the impact of some of these initiatives. The Appeals Council takes about a year to process a claimant's request for review.³¹⁰ A claimant's case typically sits for more than a year at a district court before it gets decided.³¹¹ An ALJ's decision rendered sometime in early 2013, in other words, might not cross a magistrate judge's desk until mid-2015. The ALJ almost certainly would not have used the eBB to help render the decision under review. The agency had completed fifty focused reviews by mid-2014;³¹² an ALJ counseled in 2014 would not have a decision before a federal judge until 2016. The Appeals

business cycle. However, since N_0 would have been larger in such a year (and since a_0 and D_0 are unaffected by the economy) the number of appealed cases that are meritless, A_0 , would be greater than usual. As a result, the set of cases appealed to the Appeals Council would have a greater-than-usual share that are meritless. Now let the Appeals Council remand rate be r_0 for meritless appeals and r_H for high quality appeals. In that case, the Appeals Council's remand rate is $[r_H A_H + r_0 A_0] / (A_H + A_0)$, which is lower when A_0 is greater. Assuming that the Appeals Council's remand decision for a claim is based only on claim quality, and not on the state of the business cycle when the claim was filed, the parameters r_H and r_0 can be treated as fixed with respect to the economy. It can then be shown that an increase in the number of meritless cases appealed to the Appeals Council, A_0 , will lead to a reduction in the Council's remand rate so long as r_H exceeds r_0 , which is to say, so long as the Council is more likely to remand a high quality case than a meritless one. The end result of this analysis is that, under the assumptions entertained above, the Appeals Council's remand rate will tend to be lower for a cohort of cases that has a greater share of meritless initial claims due to economic conditions, all else equal.

³⁰⁹ OIG, LOW-ALLOWANCE ADMINISTRATIVE LAW JUDGES, *supra* note 302 at 7.

³¹⁰ OIG, REQUEST FOR REVIEW WORKLOADS, *supra* note 146 at 5.

³¹¹ OFFICE OF THE INSPECTOR GENERAL, SOCIAL SECURITY ADMINISTRATION, AUDIT REPORT, OVERALL PROCESSING TIMES FOR 2010 CHILDHOOD SUPPLEMENTAL SECURITY INCOME CLAIMS, A-04-12-11230, Feb. 2013, at 11; OFFICE OF THE INSPECTOR GENERAL, SOCIAL SECURITY ADMINISTRATION, AUDIT REPORT, OVERALL DISABILITY CLAIMS TIMES FOR 2009, A-01-10-10168, May 2011, at 9.

³¹² U.S. House of Representatives, Committee on Oversight and Government Reform, *Misplaced Priorities: How the Social Security Administration Sacrificed Quality for Quantity in the Disability Determination Process*, Staff Report, 113th Cong., Dec. 18, 2014, at 12.

Council's use of clustering analysis and natural language processing did not commence until 2015.

A stagnant remand rate could possibly mask a more robust record of agency success in the federal courts. As noted, ALJ allowance rates have declined significantly, while Appeals Council agree rates have increased.³¹³ The agency maintains that this change has resulted from an improvement in ALJ performance. Assuming that this explanation is accurate, it still leaves a larger set of potential appeals from which claimant representatives can pick their cases. The ALJ allowance rate has dropped from 70% to 45%. Even if ALJs now correctly deny four out of every five claims that would have been allowed previously, the 20% error rate could increase the number of flawed ALJ decisions to challenge by thousands.³¹⁴ If the Appeals Council does not catch all of these errors, and if the private social security bar cannot expand to keep pace with the increased number of potential appeals,³¹⁵ then the strength of each appeal to the federal courts, on average, could increase. If federal judges are applying the same standard of review as they were applying, say, five years earlier, and assuming that the entire pool of claims has not changed significantly in its composition, then the remand rate should be higher than 45%. A stagnant or even modestly dropping rate, in other words, might be consistent with a more deferential federal judiciary.

³¹³ See *infra* notes 301 & 302 & accompanying text.

³¹⁴ The point here is not that any given claim's strength is affected by the agency's quality assurance initiatives. Rather, it is that there will be a larger number of incorrectly denied high quality claims in the pool from which claimant representatives—lawyers—can pick and choose in deciding whom to represent. If lawyers are able to do so with at least some efficacy, then the average quality of cases appealed to the district courts will be greater following agency quality initiatives than before.

³¹⁵ The number of potentially appealable agency decisions has increased faster than the rate of filings in the federal courts. The Appeals Council denied review in 76,605 cases in FY 2010 and 136,113 in FY 2014, for an increase of 44%. Claimants filed 12,256 cases in federal court in FY 2010 and 18,503 in FY 2014, for an increase of 34%. See www.ssa.gov/appeals/DataSets/AC02_AC_GrantReview_All_Dispositions; www.ssa.gov/appeals/DataSets/AC-5_Court_Remands_NCC_Filed.

Lawyers' incentives, a related phenomenon, give further reason to think that this enlarged pool is generating stronger appeals on average.³¹⁶ Claimant representatives typically work on a contingency fee basis while the case is pending before the agency, pegged to the amount of back benefits a successful claimant wins.³¹⁷ However, when a case goes to federal court and the court remands a case for further proceedings, a claimant representative may get paid fees through the Equal Access to Justice Act ("EAJA") for the representation she afforded in federal court.³¹⁸ If the claimant then prevails on remand and receives back benefits, then the claimant representative may also receive payment from the contingency fee.³¹⁹ To obtain all possible fees, therefore, a claimant representative must have a client prevail not only before the district court, but also before the ALJ on remand.³²⁰ The opportunity cost of federal court litigation adds to this risk of nonpayment and may further refine the metrics lawyers use to select appeals.³²¹ Given this fee structure, one claimant representative told us, she and her colleagues "are not going to file a bunch of garbage in federal court."³²²

³¹⁶ Several agency officials made this observation in conversations with us. OGC Lawyer 4 at 4 (claiming that, as an OGC lawyer defending ALJ decisions in federal court, he sees the "worst" "5-10%" of ALJ decisions); Agency Official 3, First Interview at 5 (observing that district courts see only a "tiny percentage" of ALJ decisions, and that "most are the worst cases" due to attorney incentives).

³¹⁷ Attorney compensation in social security appeals involves a complicated intersection of legal regimes, but the basic point about contingency is simple. *See generally* *Gisbrecht v. Barnhart*, 535 U.S. 789 (2002) (discussing fee arrangements in social security representation).

³¹⁸ *Cf. Shalala v. Schaefer*, 509 U.S. 292, 302 (1993) (holding that a party obtaining a remand for further proceedings in certain instances counts as a "prevailing party" for EAJA purposes).

³¹⁹ For an example of these two stages, see *Ibrahim v. Colvin*, Civ. No. 13-65, 2015 WL 4507523 (E.D. Cal. July 24, 2015).

³²⁰ *Cf.* Claimant Representative 4 at 3 (indicating that she accounts for the likelihood of prevailing before both the district court and the agency on remand when deciding which cases to appeal to the district court); Claimant Representative 12 at 2 (same). Some attorneys write briefs for other claimant representatives. They may get paid by the hour by the claimant representative who hires them. Claimant Representative 10 at 1. Regardless, attorneys' incentives are similar to what we describe above. EAJA fees require that the claimant succeed on appeal. Presumably a claimant representative's resources for hiring brief writers come from fees earned when claimants prevail.

³²¹ Claimant Representative 9 at 3. During the 40-50 hours one claimant representative allots to a federal court appeal, he told us, he could represent three or four other people at the ALJ level. Claimant Representative 1 at 1.

³²² Claimant Representative 10 at 1; *see also* Claimant Representative 9 at 3 (insisting that claimant representatives are generally selective about their case selection, but acknowledging that some lawyers are "terrible" and will bring cases even though "they are losing 90%").

At least superficially the numbers support this story. Claimants take only about 15% of appealable Appeals Council decisions to the federal courts.³²³ This case selectivity comes after another layer of selection. Claimants take only slightly more than half of ALJ decisions denying benefits to the Appeals Council. A pool composed at least in some significant measure of the weakest of ALJ decisions is the likely result. Coupled with the growing pool of potentially appealable decisions, attorney incentives should ensure a strong set of appeals to the federal courts, a situation that explains a robust remand rate.

2. The Substantial Evidence Standard of Review

SSA officials gave us another reason to regard the 45% remand rate as excessive. Perhaps federal courts often disregard the substantial evidence standard of review that Congress set for the review of ALJ decisions.³²⁴ If so, courts assume more decision-making prerogative than applicable law allows. Several prominent commentators have expressed this sentiment,³²⁵ and agency officials have voiced it for decades.³²⁶

Various formulations pepper federal courts' discussion of the substantial evidence standard, but a recent Fifth Circuit articulation gets at the classic version: "Substantial evidence is more than a scintilla, less than a preponderance, and is such relevant evidence as a reasonable mind might accept as adequate to support a conclusion."³²⁷ A court looks at the record as a

³²³ Compare www.ssa.gov/appeals/DataSets/AC02_AC_GrantReview_All_Dispositions, with www.ssa.gov/appeals/DataSets/AC-5_Court_Remands_NCC_Filed.

³²⁴ 42 U.S.C. § 405(g).

³²⁵ E.g., Verkuil, *Outcomes Analysis*, *supra* note 297 at 707-709; Verkuil & Lubbers, *supra* note 23 at 741; ACUS, TREATING PHYSICIAN RULE, *supra* note 42 at 23. Cf. CAROLYN A. KUBITSHEK & JON C. DUBIN, SOCIAL SECURITY DISABILITY: LAW AND PROCEDURE IN FEDERAL COURT § 8:2, at 991 (2014) (acknowledging that "[c]ircuits which tend to be sympathetic to claimants have found ways to stretch the narrow substantial evidence standard in order to provide what is in fact a broader scope of review").

³²⁶ SUSAN GLUCK MEZEY, NO LONGER DISABLED: THE FEDERAL COURTS AND THE POLITICS OF SOCIAL SECURITY 55 (1988).

³²⁷ Sun v. Colvin, --- F.3d ---, 2015 WL 4393795, at *5 (5th Cir. July 17, 2015) (internal quotation omitted). Compare *Universal Camera Corp. v. NLRB*, 340 U.S. 474, 477 (1951) (describing "substantial evidence" as "more than a mere scintilla. It means such relevant evidence as a reasonable mind might accept as adequate to support a conclusion.").

whole, including evidence supporting and opposing the agency’s decision, to determine if this standard is met.³²⁸

In theory, this “very deferential” standard imposes an “extremely limited” role on federal courts when they review ALJ factual findings.³²⁹ In reality, many OGC lawyers and other agency personnel told us, federal courts ignore the standard of review and often adjudicate factual questions de novo.³³⁰ If federal courts do so, then they usurp decision-making power that Congress has allocated to the SSA, and claimants win too many cases.³³¹ In contrast, most of the federal judges we interviewed pleaded “not guilty” to the agency’s charge.³³²

Whether and how federal judges stray from the deference the substantial evidence standard ostensibly compels are difficult questions. Courts clearly overstep their bounds when they simply reweigh the evidence or draw inferences from the evidence that differ from those reasonably drawn by ALJs. Beyond these straightforward errors, the substantial evidence issue gets ensnarled in complicated legal issues, such as whether courts may permissibly create presumptions for recurring evidentiary situations, or whether alleged errors are better characterized as ones of fact or law. Without the definitive resolution of such issues, we cannot conclude that courts routinely exceed their authority under the substantial evidence standard.

³²⁸ *Universal Camera*, 340 U.S. at 477.

³²⁹ *Elder v. Astrue*, 529 F.3d 408, 413 (7th Cir. 2008); *see also Dyrda v. Colvin*, 47 F. Supp. 3d 318, 323 (M.D.N.C. 2014); KUBITSHEK & DUBIN, *supra* note 325 at 989 (describing the substantial evidence standard of review as “highly deferential”). *But cf.* 3 CHARLES H. KOCH, JR. & RICHARD MURPHY, *ADMINISTRATIVE LAW & PRACTICE* § 9:24 (3d ed. 2010) (insisting that the substantial evidence standard “demands that the probabilities that the agency is correct be relatively high”); RICHARD A. POSNER, *DIVERGENT PATHS: THE ACADEMY AND THE JUDICIARY* 152-153 (2016) (defending a refusal to afford significant deference to ALJ determinations).

³³⁰ One OGC lawyer explained that judicial disregard of the substantial evidence standard is one of the two biggest problems the agency faces when litigating in federal court. OGC Lawyer 22 at 1; *see also* OGC Lawyer 21 at 1; OGC Lawyer 16, First Interview at 2; OGC Lawyer 10 at 4; OGC Lawyer 5, Second Interview at 1; OGC Lawyer 11 at 1; OGC Lawyer 3 at 5; OGC Lawyer 12 at 4; OGC Lawyer 2 at 3; OC Lawyer 13 at 4.

³³¹ *See Verkuil & Lubbers*, *supra* note 23 at 741 (interpreting the remand rate in this manner).

³³² Federal Judge 15 at 3; *see also* Federal Judge 13 at 2; Federal Judge 3 at 2; Federal Judge 7 at 1; Federal Judge 8 at 2; Federal Judge 20 at 4; Federal Judge 1 at 4; Federal Judge 11 at 3; Federal Judge 16 at 3; Federal Judge 19 at 6; Federal Judge 10 at 2; Federal Judge 17 at 1. Claimant representatives likewise dismissed the agency’s complaint as unfounded. *E.g.*, Claimant Representative 2 at 3-4; Claimant Representative 9 at 7. In contrast, one judge insisted that his colleagues routinely trespassed the limits that the standard of review imposes. Federal Judge 5 at 2.

a. Straightforward Errors

The substantial evidence standard has the most clear-cut significance for federal court review when the claimant challenges an ALJ's factual finding, and the record evidence related to this finding is mixed. If courts do not defer under these circumstances, as they usually should,³³³ then the standard loses its force. Thus, a magistrate judge clearly erred when he disagreed with an ALJ's determination that a claimant's carpal tunnel syndrome was not a severe impairment. He cited evidence in the record inconsistent with the ALJ's determination. But he did not fault or even mention the evidence the ALJ had relied upon and instead simply reweighed the evidence himself.³³⁴

The standard also clearly requires deference for factual inferences drawn from findings that are supported by substantial evidence.³³⁵ For this reason, a district judge erred when the judge faulted the ALJ for a residual functional capacity finding based on what the judge insisted was a faulty interpretation of a treating physician's report.³³⁶ Conceding that the report was "not clear," the judge drew what she thought was a more plausible inference from the report than what the ALJ had inferred.³³⁷

Finally, a court clearly errs if it faults an ALJ for failing to address inconsistencies in medical source opinions and records, if a reasonable person could conclude that the evidence is not inconsistent. Thus, a magistrate judge exceeded her authority when she faulted the ALJ for failing to give reasons to privilege one physician's opinion over another, when the opinions, fairly read, did not necessarily conflict with each other.³³⁸

³³³ Howard v. Massanari, 255 F.3d 577, 584 (8th Cir. 2001).

³³⁴ Parent v. Colvin, Civ. No. 14-3056, 2015 WL 1564886, at *4 (W.D. Ark. Apr. 8, 2015).

³³⁵ E.g., Schaudeck v. Comm'r of Social Security, 181 F.3d 429, 431 (3d Cir. 1999); Young v. Apfel, 221 F.3d 1065, 1068 (8th Cir. 2001).

³³⁶ Hiller v. Astrue, Civ. No. 11-4131, 2012 WL 4511374, at *2 (E.D.N.Y. Sept. 28, 2012).

³³⁷ *Id.* at *2.

³³⁸ Nicholson v. Astrue, Civ. No. 10-2010, 2012 U.S. Dist. LEXIS 146069, at *23-24 (D.D.C. Mar. 14, 2012).

b. Reason Giving and the Substantial Evidence Standard

Based on our interviews, we have identified at least three issues significantly more complicated than these straightforward errors that confound any easy conclusion that district courts routinely trespass limits that the substantial evidence standard sets. The first involves the relationship between the standard and reason-giving requirements prescribed in statutory text or a regulation. The treating source regulation provides an illustration. It requires the ALJ to give “controlling weight” to a treating source’s opinion on the nature and severity of an impairment if it “is well-supported by medically acceptable clinical and laboratory diagnostic techniques and is not inconsistent with the other substantial evidence” in the record. If the ALJ does not afford the opinion such weight, the ALJ must “give good reasons” for the decision.³³⁹ The Ninth Circuit reads this language as a license to scrutinize the ALJ’s decision carefully:

Where the treating doctor’s opinion is not contradicted by another doctor, it may be rejected only for clear and convincing reasons supported by substantial evidence in the record Even if the treating doctor’s opinion is contradicted by another doctor, the ALJ may not reject this opinion without providing specific and legitimate reasons supported by substantial evidence in the record This can be done by setting out a detailed and thorough summary of the facts and conflicting clinical evidence, stating his interpretation thereof, and making findings The ALJ must do more than offer his conclusions. He must set forth his own interpretations and explain why they, rather than the doctors’, are correct.³⁴⁰

Other circuits have provided similar elaborations.³⁴¹

On one hand, such formulations seem to depart dramatically from the deference the substantial evidence standard commands. How can a court require “clear and convincing” reasons for the ALJ’s decision, while still reviewing the decision deferentially?³⁴² Indeed, courts

³³⁹ 20 C.F.R. §§ 404.1527(c)(2), 416.927(c)(2).

³⁴⁰ *Reddick v. Chater*, 157 F.3d 715, 725 (9th Cir. 1998) (internal quotations and citations omitted).

³⁴¹ *E.g.*, *Burgess v. Astrue*, 537 F.3d 117, 129 (2d Cir. 2008); 2 SARAH H. BOHR ET AL., *BOHR’S SOCIAL SECURITY ISSUES ANNOTATED* § 202.8 (2012).

³⁴² ACUS, *ASSESSING THE EFFICACY*, *supra* note 42 at 23 (suggesting that the treating physician rule “has distorted substantial evidence in the record review”).

reverse and remand the agency when the medical evidence is conflicting, ordinarily a situation for deference,³⁴³ when reasons for disagreeing with the treating physician are not sufficiently compelling.³⁴⁴ On the other hand, if courts must afford the same deference to the ALJ's treating source findings as to other findings of fact, then courts lose their power to ensure that ALJs comply with the "good reasons" requirement the regulation provides.

To our knowledge, the federal courts have not explicitly analyzed the significance a heightened reason-giving requirement has for the deference otherwise owed under the substantial evidence standard, at least in these terms. The proper way to understand this relationship is thus unresolved and begs for clarification, whether through appellate engagement, rulemaking, or even legislation. Indeed, this sort of relationship is contested more generally. In immigration cases, the circuits have split over how burdens of proof affect substantial evidence review, an analogous problem. The Second Circuit insists that "the substantial evidence test becomes more demanding as the government's underlying burden of proof increases,"³⁴⁵ while the Eleventh Circuit denies that the agency's increased burden affects the stringency of its review.³⁴⁶

We recommend one way to approach the relationship between the substantial evidence standard and a heightened reason-giving obligation. Their blurry relationship can come into somewhat better focus if one appreciates that the standard not only governs review of the evidentiary basis for the ALJ's decision, but also review of the reasons the ALJ gives connecting the evidence to the conclusions drawn.³⁴⁷ Assume that a treating source opines that two claimants have severe, debilitating ankle pain. The ALJ #1 rejects this opinion, based in part on a finding that Claimant #1 runs five miles per day. The evidence of this exercise is mixed,

³⁴³ *E.g.*, Wright v. Massanari, 321 F.3d 611, 614 (6th Cir. 2003).

³⁴⁴ *E.g.*, Komiyama v. Astrue, Civ. No. 09-159, 2009 WL 4730586, at *2-*3 (C.D. Cal. Dec. 4, 2009).

³⁴⁵ Francis v. Gonzales, 442 F.3d 131, 138 (2d Cir. 2006).

³⁴⁶ Adefemi v. Ashcroft, 386 F.3d 1022, 1028 (11th Cir. 2004).

³⁴⁷ 3 PIERCE, *supra* note 11 § 11.2, at 988; Carpenter & Millwrights, Local Union 2471 v. N.L.R.B., 481 F.3d 804, 809 (D.C. Cir. 2007).

Claimant #1 disputes it, and a reasonable person could both agree and disagree with ALJ #1's finding. ALJ #2 rejects this opinion, based in part on a finding that Claimant #2 performs household chores and takes care of her child. The evidence of these activities is likewise mixed.

The two cases are identical in terms of the equivocal evidence supporting the ALJs' reasons. But in the first case, the ALJ's reason itself is clearly a good one. The court should not take the reason-giving obligation as an invitation to engage in a *de novo* review of the ALJ's evidentiary finding with regard to the exercise.³⁴⁸ In the second case, it is not the equivocal evidence but rather the insufficiently good reason that justifies a decision in the claimant's favor.³⁴⁹ The substantial evidence standard commands deference by requiring the reviewing court to allow the ALJ's evidentiary finding with regard to the activities to stand, but the reason-giving obligation allows the court to demand a more plausible reason, based on the evidence, to support the conclusion drawn.³⁵⁰

In many instances, the reason the ALJ gives does not cleave so neatly from the underlying evidence in the record. An ALJ might, for example, reject a treating source's opinion as to an impairment's severity on grounds that it is inconsistent with other medical evidence in the record. Under such circumstances, courts may insist that the ALJ provide specific reasons for his determination, but the court, again, should resist the temptation to take the reason-giving obligation as an invitation for *de novo* review. The court's job is to ensure that the ALJ considered the evidence thoughtfully, not that he made the right evidentiary determination.

³⁴⁸ The Sixth Circuit describes the "good reasons" requirement in the treating source rule as a "procedural requirement" that "denotes a lack of substantial evidence, even where the conclusion of the ALJ may be justified based upon the record." *Rogers v. Comm'r of Social Security*, 486 F.3d 234, 243 (6th Cir. 2007). Although confusing, this statement suggests that the failure to discharge a "procedural requirement" means that the ALJ's decision does not benefit from deference. The form of review is equivalent to that for legal error, that is, *de novo*. *See also* *Lingenfelter v. Astrue*, 504 F.3d 1028, 1038 n.10 (9th Cir. 2007).

³⁴⁹ *Cf.* *Hughes v. Astrue*, 705 F.3d 276, 278 (7th Cir. 2013) (Posner, J.) (criticizing ALJs for relying on evidence of household activities to support findings of no disability).

³⁵⁰ *E.g.*, *Whitney v. SEC*, 604 F.2d 676, 681 (D.C. Cir. 1979).

c. Judicial Elaborations

A second confounding issue involves adjustments that courts have fashioned to the substantial evidence standard, without being prompted to do so by a reason-giving requirement or something like it set forth in positive law. The Ninth Circuit, for example, insists that an ALJ must give “clear and convincing reasons” to reject the claimant’s description of his symptoms “where the record includes objective medical evidence establishing that the claimant suffers from an impairment that could reasonably produce the symptoms of which he complains”³⁵¹ This threshold has no obvious basis in either the Act or applicable regulations and seems to have originated in case law.³⁵²

This “clear and convincing” threshold, the Ninth Circuit insists, “is not an easy requirement to meet.”³⁵³ For this reason the agency believes that it conflicts with the deferential substantial evidence standard.³⁵⁴ A couple of other courts have used language similar to the Ninth Circuit’s,³⁵⁵ but the “clear and convincing” threshold has hardly won a universal embrace.³⁵⁶ Moreover, the threshold has produced garbled recitations of the substantial evidence standard, suggesting that it has sown confusion over the appropriate degree of deference courts owe.³⁵⁷

³⁵¹ *Carmickle v. Comm’r of Social Security*, 533 F.3d 1155, 1160 (9th Cir. 2008); *see also* *Santiago v. Astrue*, Civ. No. 06-302, 2010 WL 466052, at *13 (D. Ariz. Feb. 10, 2010).

³⁵² *E.g.*, *Larson v. Colvin*, Civ. No. 12-1844, 2013 WL 3216172, at *6 n.1 (D. Ariz. June 25, 2013). We traced the “clear and convincing” standard back to *Day v. Weinberger*, 522 F.2d 1154, 1156 (9th Cir. 1975), which in turn imported it from *White Glove Building Maintenance, Inc. v. Brennan*, 518 F.2d 1271 (9th Cir. 1971), a case involving an appeal from a Department of Labor decision.

³⁵³ *Garrison v. Colvin*, 759 F.3d 995, 1015 (9th Cir. 2014).

³⁵⁴ The agency argues that this is so. *E.g.*, *Webber v. Colvin*, Civ. No. 13-2088, 2015 WL 1810205, at *8 (D. Ore. Apr. 17, 2015).

³⁵⁵ *Schaudeck v. Comm’r of Social Security*, 181 F.3d 429, 433 (3d Cir. 1999) (insisting that the ALJ must give symptom complaints “great weight” when consistent with objective medical evidence); *Hersch v. Barnhart*, 470 F. Supp. 2d 1281, 1284 (D. Utah 2006) (same).

³⁵⁶ *Smith v. Astrue*, 457 Fed. Appx. 326, 329 (4th Cir. 2011) (rejecting a “great weight rule”).

³⁵⁷ The Ninth Circuit, for instance, once asked “whether the ALJ’s adverse credibility finding . . . is supported by substantial evidence under the clear-and-convincing standard.” *Carmickle*, 533 F.3d at 1161.

However problematic, the Ninth Circuit’s elaboration and ones like it are not necessarily illegitimate.³⁵⁸ The “clear and convincing” threshold could be viewed as a power grab at odds with a statutory insistence on deference. Viewed differently, however, the threshold may be an attempt to inject some objectivity into the “seedless grape” of substantial evidence review.³⁵⁹ As two commentators aptly noted, “it is difficult to ‘consider’ the evidence contrary to the agency’s finding, which is required, without reweighing the evidence, which the reviewing court is forbidden from doing.”³⁶⁰ Without stated rules for its application in recurring circumstances, the substantial evidence standard invites inconsistent decision-making. The Ninth Circuit’s threshold might be seen as an attempt to give more direction to lower courts’ review, one with a certain logic behind it. The substantial evidence standard is the same – what could a reasonable person conclude? But a reasonable person would be less likely to conclude that the claimant is exaggerating or lying about her symptoms, so the logic goes, if they are consistent with objective medical evidence. Hence, the reasonable person would require clear, specific reasons before disbelieving the claimant.

Interpreted thusly, the “clear and convincing” threshold is not a judicial trump but a presumption guiding the proper application of the substantial evidence standard in a recurring circumstance. If so, the concern is not one of illegitimate usurpation but of proper policy. Is the presence of objective medical evidence sufficiently correlated with claimant veracity to support the Ninth Circuit’s presumption? If so, what determination ought an ALJ make to find a claimant’s symptoms not supported or not consistent with the record, notwithstanding this evidence? The agency is better positioned than the courts to answer these questions, and courts

³⁵⁸ *But see* 3 PIERCE, *supra* note 11 § 11.3, at 1015 (questioning the legitimacy of such elaborations).

³⁵⁹ Ernest Gellhorn & Glen O. Robinson, *Perspectives on Administrative Law*, 75 COLUM. L. REV. 771, 780 (1975); *see also* CHRISTOPHER EDLEY, JR., *ADMINISTRATIVE LAW: RETHINKING JUDICIAL CONTROL OF BUREAUCRACY* 97 (1990) (suggesting that existing standards of review are in effect “lawless[.]”).

³⁶⁰ Richard E. Levy & Robert L. Glicksman, *Agency-Specific Precedents*, 89 TEX. L. REV. 499, 544 (2011).

should take care to defer to its guidance if offered. If, however, the agency has not fashioned guidance, interest in consistent decision-making may justify the judicial elaboration.³⁶¹ More generally, judicial elaborations on the substantial evidence standard should invite the agency to clarify how best to resolve the evidentiary concerns that are motivating courts to fashion these presumptions.

d. Characterization Problems

A third complication is the most difficult one and prevents any simple answer to the question of whether the federal courts are unfaithful to the substantial evidence standard. Difficulties of characterization – is an alleged error one of fact or law? – abound.³⁶² The difference matters, because a federal judge owes no deference to the ALJ when the appeal involves an alleged legal error.³⁶³ To many of our interview subjects, characterization is the whole ballgame. The agency insists that the alleged error is really a dispute of fact, while the claimant representative couches it as a legal error.³⁶⁴

Several examples illustrate unresolved characterization problems that produce less deferential review than what the substantial evidence standard would otherwise predict:

- The SSA’s interpretation of a regulation governing medical source evidence requires ALJs to evaluate all of the medical evidence in the record.³⁶⁵ To some courts, the failure

³⁶¹ A newly issued Social Security Ruling offers detailed guidance on some of the issues implicated by the Ninth Circuit’s elaboration. Soc. Sec. Ruling 16-3p.

³⁶² Cf. Zaring, *supra* note 273 at 152 (“Distinguishing between questions of law and questions of fact has never been easy either, at least at the margin.”).

³⁶³ E.g., *McClanahan v. Comm’r of Social Security*, 474 F.3d 830, 833 (6th Cir. 2006); *Poulos v. Comm’r of Social Security*, 474 F.3d 88, 91 (3d Cir. 2007); *McClellan v. Astrue*, 650 F. Supp. 2d 223, 226-27 (E.D.N.Y. 2009); *Seavey v. Barnhart*, 276 F.3d 1, 9 (1st Cir. 2001).

³⁶⁴ OGC Lawyer 4 at 4; OGC Lawyer 5, Second Interview at 2. A number of claimant representatives insisted to us that they rarely seek review of ALJ findings of fact and instead challenge decisions for legal error. Claimant Representative 8 at 4; Claimant Representative 2 at 3; Claimant Representative 10 at 4; Claimant Representative 7 at 4; Claimant Representative 1 at 5; Claimant Representative 9 at 7; Claimant Representative 6 at 3; Claimant Representative 7 at 4 (“Deference doesn’t have anything to do with it.”); *see also* Federal Judge 3 at 2 (explaining that remands happen because ALJs don’t follow “regulations and the law”).

³⁶⁵ 20 C.F.R. §§ 404.1527(c), 416.1527(c); *see also* Soc. Sec. Ruling 96-5p (“[O]pinions from any medical source on issues reserved to the Commissioner must never be ignored. The adjudicator is required to evaluate all evidence in the case record that may have a bearing on the determination or decision of disability, including opinions from medical sources about issues reserved to the Commissioner.”).

to discuss a medical opinion in the record is “clear legal error” and requires a remand unless the error is harmless.³⁶⁶ This harmless error analysis then flips the substantial evidence standard on its head: the district court must remand unless no reasonable ALJ could conclude that the overlooked medical opinion could change the disability determination.³⁶⁷ For other courts, an ALJ conceivably could omit mention of a particular opinion and still have his decision supported by substantial evidence.³⁶⁸

- To obtain benefits, a claimant seeking SSDI benefits must establish the onset of a disability before his “date last insured.”³⁶⁹ To some courts, an ALJ must use a medical advisor to help determine the onset date if the available evidence is ambiguous. The failure to do so is reversible legal error.³⁷⁰ To others, the ALJ has no such obligation, and thus whether the ALJ rightly determined the onset date is reviewed for substantial evidence.³⁷¹
- The circuits are split on the issue of whether an ALJ must make specific findings to support a determination that a claimant has transferable skills and thus is not disabled, even if she cannot continue in past employment. The Sixth Circuit has disclaimed any specific findings requirement when the ALJ relies on a vocational expert, and when the expert testifies to the claimant’s transferable skills.³⁷² The determination of transferable skills is reviewed for substantial evidence. In the Ninth Circuit, the ALJ must make the findings, and failure to do so is legal error.³⁷³
- To some courts, an ALJ who makes an RFC determination based on “raw medical evidence” and without a medical source opinion as to the claimant’s capacity commits a legal error. To others, the ALJ’s RFC determination simply needs to be supported by substantial evidence, whether it is based on a medical source opinion or not.³⁷⁴

An important unresolved legal problem involves the *Chenery* Doctrine’s relevance. The agency has argued that substantial evidence review is supposed to be based on anything in the record that arguably supports the ALJ’s decision, and not just on the evidence the ALJ actually

³⁶⁶ Sandoval v. Colvin, Civ. No. 14-1852, 2015 WL 4245824, at *2 (D. Colo. July 14, 2015).

³⁶⁷ Fischer-Ross v. Barnhart, 431 F.3d 729, 733-34 (10th Cir. 2005); *see also* Spiva v. Astrue, 628 F.3d 346, 353 (7th Cir. 2010); Neal ex rel. Walker v. Barnhart, 405 F.3d 685, 689 (8th Cir. 2005); Fitzgerald v. Astrue, Civ. No. 08-170, 2009 WL 4571762, at *9 (D. Vt. Nov. 30, 2009); Davis v. Astrue, Civ. No. 07-231, 2008 WL 540899, at *3 (D.S.C. Feb. 22, 2008).

³⁶⁸ *E.g.*, Thompson v. Colvin, Civ. No. 12-585, 2013 WL 1718768, at *14 (N.D. Ill. Apr. 18, 2013).

³⁶⁹ Soc. Sec. Ruling 83-20.

³⁷⁰ Wilson v. Colvin, 17 F. Supp. 3d 128, 139 (D.N.H. 2014).

³⁷¹ Eichstadt v. Astrue, 534 F.3d 663, 667 (7th Cir. 2008).

³⁷² Wilson v. Comm’r of Social Security, 378 F.3d 541, 549-50 (6th Cir. 2004).

³⁷³ Bray v. Comm’r of Social Security, 554 F.3d 1219, 1225 (9th Cir. 2009); Coyne v. Astrue, Civ. No. 11-868, 2012 WL 1698318, at *4 (C.D. Cal. May 14, 2012).

³⁷⁴ *See generally* Poitra v. Colvin, Civ. No. 14-094, 2015 WL 9700337, at *24-29 (D.N.D. Sept. 30, 2015).

marshaled in her opinion.³⁷⁵ But the *Chenery* Doctrine provides that “agency action can be upheld, if at all, only on the rationale the agency itself articulated when taking action.”³⁷⁶ Representing the majority view,³⁷⁷ the Seventh Circuit has rejected the SSA’s argument and insists that *Chenery* precludes the SSA from offering post-hoc rationalizations for an ALJ’s decision.³⁷⁸ The substantial evidence calculus cannot include any evidence not discussed by the ALJ, however compelling. In contrast, some courts suggest, either implicitly³⁷⁹ or explicitly,³⁸⁰ that *Chenery* does not apply in social security cases. They agree that the question on review is “whether the ALJ’s decision is supported by substantial evidence based upon the record taken as a whole,” regardless of what the ALJ actually discusses.³⁸¹ Other courts have staked out a third position: an ALJ does not need to discuss evidence for a district court to consider it in the substantial evidence balance, so long as “the [ALJ’s] decision . . . demonstrate[s] that the ALJ considered all of the evidence”³⁸²

³⁷⁵ OGC Lawyer 14 at 4; *see also* Federal Judge 10 at 2 (observing that he as a federal judge often finds himself saying, “if I were reviewing the Commissioner’s brief, there’s substantial evidence”); Claimant Representative 1 at 5; Claimant Representative 9 at 6 (insisting as a claimant representative that the typical OGC brief has “5-15” post-hoc rationalizations); *Spiva v. Astrue*, 628 F.3d 346, 353 (7th Cir. 2010) (commenting on a recurrent theme in government briefs); *Slade v. Colvin*, Civ. No. 12-1253, 2014 WL 580895, at *3 (E.D. Cal. Feb. 13, 2014); *Nimmerrichter v. Colvin*, 4 F. Supp. 3d 958, 971 (N.D. Ill. 2013); *Gibson v. Astrue*, Civ. No. 09-677, 2010 WL 3655857, at *10 (N.D. Ga. Sept. 13, 2010).

³⁷⁶ Elizabeth Magill & Adrian Vermeule, *Allocating Power Within Agencies*, 120 YALE L.J. 1032, 1042 (2011) (citing *SEC v. Chenery Corp.*, 318 U.S. 80, 94 (1943)).

³⁷⁷ Bryan C. Bond, Note, *Taking it on the Chenery: Should the Principles of Chenery I Apply in Social Security Disability Cases?*, 86 NOTRE DAME L. REV. 2157, 2159 & n.11 (2011).

³⁷⁸ *E.g.*, *Hanson v. Colvin*, 760 F.3d 759, 761 (7th Cir. 2014) (suggesting that *Chenery* violations are a “recurrent feature” of SSA arguments, and that they amount to “professional misconduct” that might warrant sanctions); *Spiva v. Astrue*, 628 F.3d 346, 353 (7th Cir. 2010) (criticizing the SSA’s litigation position as “determined to dissolve the *Chenery* doctrine in an acid of harmless error”); *Parker v. Astrue*, 597 F.3d 920, 922 (7th Cir. 2010).

³⁷⁹ *E.g.*, *Rector v. Colvin*, Civ. No. 14-5017, 2015 WL 1925368, at *3-*4 (W.D. Ark. Apr. 28, 2015) (undertaking an independent survey of the record evidence to find substantial evidence in support of the ALJ’s decision).

³⁸⁰ District courts find sufficient ambiguity in the Sixth Circuit’s treatment of the *Chenery* issue to conclude that they are not obliged to apply the *Chenery* doctrine in social security cases. *Calvin v. Astrue*, Civ. No. 05-0109, 2010 WL 55452, at *14 (M.D. Tenn. Jan. 7, 2010); *Brasseur v. Commissioner of Social Security*, Civ. No. 10-1260, 2012 WL 991621, at *6 (W.D. Mich. Feb. 3, 2012). The Tenth Circuit has been ambiguous as well. *See* Bond, *supra* note 377 at 2159 n.12. For dicta denouncing the Seventh Circuit’s application of *Chenery* to the social security context, *see Senn v. Astrue*, Civ. No. 12-326, 2013 WL 639257, at *6-9 (W.D. Wis. Feb. 21, 2013).

³⁸¹ *Brasseur v. Commissioner of Social Security*, Civ. No. 10-1260, 2012 WL 991621, at *6 (W.D. Mich. Feb. 3, 2012).

³⁸² *Buchan v. Astrue*, Civ. No. 10-4081, 2011 WL 3714472, at *5 (D. Kan. Aug. 24, 2011).

The 45% remand rate indicates widespread disregard of the substantial evidence standard of review only if the claimant-friendly characterizations of these issues are pre-textual (i.e., made to cloak a review of factual findings) or manifestly wrong. Some probably are.³⁸³ But other issues, such as the *Chenery* Doctrine's application, are complicated and contestable.³⁸⁴ If a federal judge can defensibly characterize certain shortcomings as legal errors, then the appropriate scope of application of the substantial evidence standard is debatable.

As before, the agency could take asserted legal errors as an invitation for rulemaking or the issuance of a ruling. The obligation that some courts have put on ALJs to obtain medical source statements on the claimant's capacity, for example, seems to have originated in case law. A regulation or ruling indicating when or under what circumstances an ALJ may eschew a medical source statement could lay this doctrine to rest and reinvigorate the substantial evidence standard's deference for such issues.

3. The Missed Opportunity of Appellate Review

While we do not wade into the merits of any particular characterization dispute, two phenomena do deserve criticism. Sometimes federal judges persist with articulation or record development requirements when either circuit case law or the applicable regulation plainly does

³⁸³ A recently superseded ruling identified seven factors that the agency "will consider" when evaluating the credibility of a claimant who alleges symptoms that lack support in objective medical evidence. Soc. Sec. Ruling 96-7p, *superseded by* Soc. Sec. Ruling 16-3p. By one interpretation of this regulation, an ALJ could consider only some of the factors and still make a credibility determination supported by substantial evidence. *E.g.*, *Routh v. Astrue*, 698 F. Supp. 2d 1072, 1078 (E.D. Ark. 2010) (declaring that an ALJ does not have to discuss each of the credibility factors explicitly); *id.* at 1081 (finding the ALJ's credibility determination supported by substantial evidence); *see also* *Michael v. Astrue*, Civ. No. 09-321, 2010 WL 2868177, at *5 (N.D. Ohio June 28, 2010). To other courts, an ALJ's failure to discuss each of the seven factors was legal error that warranted a remand, even if the reasons the ALJ did provide might otherwise provide substantial evidence in support of a credibility determination. *E.g.*, *Henningsen v. Comm'r of Social Security*, --- F. Supp. 3d ---, Civ. No. 13-4392, 2015 WL 3604912, at *16 (E.D.N.Y. June 8, 2015); *Taylor v. Comm'r of Social Security*, Civ. No. 13-5995, 2014 WL 2465057, at *13 (S.D.N.Y. May 21, 2014). The latter position was probably incorrect. *E.g.*, *Chichoki v. Astrue*, 534 Fed. Appx. 71, 76 (2d Cir. 2013) (declaring that an ALJ does not need to discuss all seven credibility factors).

³⁸⁴ *See generally* *Bond*, *supra* note 377.

not impose such obligations on ALJs.³⁸⁵ Here the agency can rightly complain of ever-shifting legal standards that its ALJ corps can never quite meet.³⁸⁶ Also, given numerous circuit splits, one wonders if the federal courts of appeals have fully appreciated the agency's interest in the administration of a national program consistently.

What to do about these problems of non-conforming judges and inter-circuit disuniformity? This question can be merged with a second: to the extent that the agency believes the federal courts have characterized issues improperly or have imposed undue requirements on ALJs, what can be done? Obviously, the agency can clarify its views through notice-and-comment rulemaking. The issuance of new interpretive rules is a less powerful but administratively easier tactic. The next obvious option for the agency is to appeal cases to the circuits in order to discipline outlier district or magistrate judges, or to obtain appellate clarification of a particular articulation or record development requirement.

Here lies a major problem for the agency. As described above, the agency must jump through an extraordinary set of hoops, culminating in Solicitor General approval, to appeal an adverse district court decision. This cumbersome process has all but shut down appeals by the agency to the circuits. In the twelve months ending December 31, 2014, 667 social security appeals were lodged with the twelve circuits.³⁸⁷ Claimants filed all but one of these.³⁸⁸ There

³⁸⁵ For instance, one federal judge faulted an ALJ's credibility finding because the ALJ did not account for the claimant's prior work history, even though "the ALJ undertook an extensive analysis of plaintiff's credibility in light of the . . . factors" actually enumerated in the applicable regulation. *Gonzalez v. Astrue*, Civ. No. 10-2941, 2012 WL 3930412, at *8 (E.D.N.Y. Sept. 10, 2012).

³⁸⁶ Some federal court standards, an ALJ complained, would require eighty hours per week of work to write a defensible decision. ALJ 2 at 2; *see also* ALJ 14 at 2 (insisting that, if this ALJ decided cases sufficient to meet a federal judge's threshold for substantial evidence, the ALJ could only hold one hearing per day).

³⁸⁷ Statistical Tables for the Federal Judiciary – December 2014, Table B-7, at 2, *at* www.uscourts.gov/statistics/table/b-7/statistical-tables-federal-judiciary/2014/12/31.

³⁸⁸ Communication with Agency Official, Nov. 18, 2015.

are good reasons to keep one's appellate powder dry, but the Department of Justice's veto has effectively given claimants complete control over the circuits' social security dockets.³⁸⁹

This dramatic asymmetry in appellate access has a number of likely effects. First, the occluded path to appellate review prevents the agency from addressing characterization concerns, correcting obvious errors, and thereby empowering the substantial evidence standard of review, if indeed the agency rightly perceives it as anemic.³⁹⁰ Presumably a claimant will not ask a court of appeals to characterize an error as one of fact, not law. To the extent that district courts treat particular issues as legal error, circuits will leave them largely alone when a claimant appeals. Second, the agency may have difficulty correcting errors or district court recalcitrance. Sometimes a district or magistrate judge might persist with a line of reasoning and argument even after the circuit has gone the other way.³⁹¹ The mistake might be honest, given the surfeit of social security case law in several circuits. It might also be a deliberate refusal to follow governing precedent. Either way, the agency's inability to appeal cuts off the appropriate avenue to respond.

Third, claimants' asymmetric access to appellate review probably produces a pro-claimant drift in circuit precedent due to claimant control over the composition of appellate dockets.³⁹² In almost every case, claimants ask appellate courts to construe governing law in a

³⁸⁹ Cf. OGC Lawyer 23 at 1 (acknowledging that the DOJ has acted as a brake); OGC Lawyer 16, First Interview at 2 (wishing that the SSA could appeal more).

³⁹⁰ We doubt that circuit or Supreme Court review could give more precise content to the substantial evidence standard itself. For one thing, as a classic critique puts it, "rules governing judicial review have no more substance at the core than a seedless grape . . ." Ernest Gellhorn & Glen O. Robinson, *Perspectives on Administrative Law*, 75 COLUM. L. REV. 771, 780 (1975); see also CHRISTOPHER EDLEY, JR., ADMINISTRATIVE LAW: RETHINKING JUDICIAL CONTROL OF BUREAUCRACY 97 (1990) (suggesting that existing standards of review are in effect "lawless[]"). For another, a circuit or the Supreme Court might balk at trying to articulate a more specific test for substantial evidence in a social security case, since any such decision would apply to all cases involving the review of agency fact-finding. The trans-substantive reality of the substantial evidence standard probably requires its articulation in ambiguous language.

³⁹¹ See *infra* note 383 (discussing credibility factors).

³⁹² See generally Jonathan Masur, *Patent Inflation*, 121 YALE L.J. 470 (2011); Melissa F. Wasserman, *Deference Asymmetries: Distortions in the Evolution of Regulatory Law*, 93 TEX. L. REV. 625, 666-667 (2015); Kate Stith, *The*

more claimant-friendly way than the district court and ALJ did, against a factual backdrop of a claimant's choosing. The numbers – about 670 claimant appeals out of nearly 10,000 claimant losses in 2014 – suggest that claimants select appeals with some care.³⁹³

This asymmetric access may also encourage pro-claimant tendencies in district and magistrate judges, even if no applicable circuit case law exists. One can safely assume that lower court judges do not like to be reversed,³⁹⁴ and findings suggest that district judges sometimes anticipate and respond to preferences of their appellate colleagues.³⁹⁵ If claimants can credibly threaten to appeal while the agency cannot, district and magistrate judges may err in favor of claimants in close cases and know that their decisions will go unreviewed.³⁹⁶

As discussed in Part IV, circuit membership explains a good deal of the inconsistencies in district court remand rates. The DOJ gauntlet the agency must run means that it lacks access to a process, appellate review, that the agency could use to try to get more uniform case law. Even more problematic is intra-circuit inconsistency, or the proliferation of conflicting standards in a

Risk of Legal Error in Criminal Cases: Some Consequences of the Asymmetry in the Right to Appeal, 57 U. CHI. L. REV. 1 (1990). One OGC lawyer told us that the Second Circuit has drifted in a pro-agency direction since the late 1990s. OGC Lawyer 5, Second Interview at 2. A claimant representative shared this impression. Claimant Representative 6 at 3. Another claimant representative thought that the Eighth Circuit has also moved in a pro-agency direction. Claimant Representative 9 at 1. A third claimant representative observed that the Seventh Circuit has moved in a pro-claimant direction over the past five years. Claimant Representative 3 at 3.

³⁹³ A couple of claimant representatives whom we interviewed complained that some of their colleagues select cases poorly for appeals. Claimant Representative 9 at 6; Claimant Representative 8 at 4. Each of those with whom we spoke insisted that they select cases to appeal very carefully. Claimant Representative 10 at 4; Claimant Representative 7 at 5; Claimant Representative 5 at 3-4; Claimant Representative 3 at 3; Claimant Representative 12 at 1. An OGC lawyer with extensive appellate experience insisted that claimants select cases very carefully. OGC Lawyer 16, First Interview at 2. *But see* OGC Lawyer 5, Second Interview at 6 (insisting that plaintiffs lawyers in the Second Circuit are not selective). The statistic in the text includes pro se appeals. We do not have sufficient data to know how many of these appellants have representation.

³⁹⁴ *E.g.*, RICHARD A. POSNER, HOW JUDGES THINK 141 (2008); Evan H. Caminker, *Precedent and Prediction: The Forward-Looking Aspects of Inferior Court Decision Making*, 73 TEX. L. REV. 1, 77 (1994); Jack Knight & Mitu Gulati, *Talking Judges*, at http://scholarship.law.duke.edu/cgi/viewcontent.cgi?article=2839&context=faculty_scholarship.

³⁹⁵ Kirk A. Randazzo, *Strategic Anticipation and the Hierarchy of Justice in the U.S. District Courts*, 36 AM. POLITICS RES. 669 (2008). *See generally* Christina L. Boyd & James F. Spriggs II, *An Examination of Strategic Anticipation of Appellate Court Preferences by Federal District Court Judges*, 29 J. LAW & POL'Y 39, 52-53 (2009) (summarizing existing literature).

³⁹⁶ Several OGC lawyers harbor this impression. OGC Lawyer 23 at 2 (claiming that district judges are aware that the agency won't appeal); OGC Lawyer 8 at 5; OGC Lawyer 11 at 3; OGC Lawyer 2 at 5; OGC Lawyer 3 at 5.

single circuit's case law. We can find a Ninth Circuit case holding whatever we need for a decision, a magistrate judge told us;³⁹⁷ by the court's own admission, this claim is not hyperbolic.³⁹⁸ The agency has many fewer opportunities to seek clarification if it effectively cannot appeal. Appellate asymmetry may also encourage renegade judges who stray from circuit guidance or the proliferation of standards that the agency believes to be unsound.³⁹⁹ Finally, asymmetry has probably engendered the proliferation of articulation and record development requirements and concomitantly blunted the force of the substantial evidence standard.

C. The Quality of Agency Adjudication

If the 45% remand rate does not necessarily reflect exaggerated scrutiny of ALJ decisions in the federal courts, then does it highlight systemic dysfunction within the agency? We do not

³⁹⁷ Federal Judge 9 at 2.

³⁹⁸ *Vasquez v. Astrue*, 572 F.3d 586, 593 (9th Cir. 2008) (discussing an intra-circuit inconsistency). Compare *Weiner v. Colvin*, Civ. No. 12-598, 2013 WL 3440021, at *5 n.1 (D. Ariz. July 9, 2013), with *Dominguez v. Colvin*, Civ. No. 12-691, 2014 WL 4904735, at *8 n.6 (D. Ariz. Sept. 23, 2014) (offering contrasting interpretations of Ninth Circuit decisions on credibility. *But see* *LoPresto v. Colvin*, Civ. No. 12-739, 2013 WL 1628727, at *7 (D. Ariz. Apr. 16, 2013) (reconciling the Ninth Circuit decisions on the issue).

³⁹⁹ A good example of the force of appellate review involves the Ninth Circuit's credit-as-true rule. When an ALJ fails to offer sufficient reasons to discount either medical opinion evidence or a claimant's credibility, the court can credit the evidence as true and remand for the calculation of benefits. *Garrison v. Colvin*, 759 F.3d 995, 1020 (9th Cir. 2014). Given a second chance to discredit opinion evidence or a claimant's testimony, an ALJ could do so adequately in some cases. In these instances, a federal court would review the ALJ's determination deferentially, for substantial evidence. Denying the agency this second bite at the apple, the credit-as-true rule likely generates benefits that would not otherwise be granted. *ACUS, TREATING PHYSICIAN RULE*, *supra* note 42 at 20; *Smith v. Astrue*, Civ. No. 11-2524, 2012 WL 5269395, at *11 n.7 (D. Ariz. Oct. 24, 2012). The rule arguably conflicts with other Ninth Circuit case law, and it has drawn criticism, including from district judges forced to apply it. *E.g.*, *Sproule v. Colvin*, Civ. No. 13-1427, 2014 WL 690988, at *4 n.2 (D. Ariz. Feb. 24, 2014); *Smith v. Astrue*, Civ. No. 11-2524, 2012 WL 5269395, at *11 n.7 (D. Ariz. Oct. 24, 2012) (Campbell, J.); *Agnew-Currie v. Astrue*, 875 F. Supp. 2d 967, 972 (D. Ariz. 2012); *Vasquez v. Astrue*, 572 F.3d 586, 601 (9th Cir. 2009) (O'Scannlain, J., dissenting); *ACUS, TREATING PHYSICIAN RULE*, *supra* note 42 at 20.

In a 2012 decision, a Massachusetts district judge relied on the Ninth Circuit's credit-as-true rule to justify an award of benefits instead of a remand for further adjudication. *Sarmiento v. Astrue*, Civ. No. 10-11724, 2012 WL 3307086, at *1 (D. Mass. Aug. 13, 2012). After the agency filed a notice of appeal with the First Circuit, the district judge granted a motion for reconsideration and vacated the part of her order incorporating the credit-as-true rule. *Sarmiento v. Astrue*, Civ. No. 10-11724, 2013 WL 427379, at *1 (D. Mass. Jan. 4, 2013); *Sarmiento v. Astrue*, Civ. No. 10-11724, 2013 WL 434184, at *1 (D. Mass. Jan. 18, 2013).

The agency has appealed to the Ninth Circuit on credit-as-true grounds, most recently in 2013. *Trnavsky v. Colvin*, --- Fed Appx. ---, 2016 WL 146007 (9th Cir. Jan. 6, 2016) (applying the credit-as-true rule). In 2009, when the Ninth Circuit muddied the credit-as-true waters, the agency attempted to get permission for a certiorari petition. But the Solicitor General said no. OGC Lawyer 23 at 1; OGC Lawyer 11 at 3. For other circuits' approaches to the issue the Ninth Circuit treats with credit-as-true, see *KUBITSHEK & DUBIN*, *supra* note 325 § 9:51 n.6. See also *Pollock v. Astrue*, Civ. No. 07-1114, 2010 WL 813522, at *1 (M.D. Fla. Mar. 3, 2010).

think so. The SSA and the federal courts differ along several institutional axes. As long as these institutions have conflicting goals and resources, honor different legal commitments, and employ decision makers with different perspectives, claimants will always win a significant number of cases in the federal courts.

1. Different Goals and Resources

On one level, the SSA and the district courts share the same goal: the accurate and efficient implementation of social security disability policy. On another, however, their goals diverge. The SSA tries to meet a quality threshold, one equaled when decisions are “factually accurate, procedurally adequate, policy compliant, and supported by the record.”⁴⁰⁰ But the agency has to serve an equally important quantity goal, the timely adjudication of huge numbers of claims.

Quality conflicts with quantity, for obvious reasons.⁴⁰¹ 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 agency is legitimately concerned with the injustice of a claim unreasonably delayed,⁴⁰² and it faces constant and enduring scrutiny for its claims backlog.⁴⁰³ Overall, the

⁴⁰⁰ Agency Official 3, Third Interview at 1.

⁴⁰¹ *Association of Administrative Law Judges v. Colvin*, 777 F.3d 402, 404-05 (7th Cir. 2015); GAO, SSA HAS TAKEN STEPS, *supra* note 208 at 4; *id.* at 17. *But see* Richard J. Pierce, *Political Control Versus Impermissible Bias in Agency Decisionmaking: Lessons from Chevron and Mistretta*, 57 U. CHI. L. REV. 481, 507 (1990) (discussing earlier agency efforts at increasing output and insisting that they did not compromise the quality of decisions). In an April 2007 report, the U.S. Government Accountability Office suggested a direct correlation between caseload pressures and quality of ALJ decision-making. Several agency officials challenged a causal link between high case loads and poor decisions, telling us that SSA had come up with several tools like the eBB to enable ALJs to dispose of a high volume of claims accurately. Agency Official 3, Third Interview at 1; OGC Lawyer 23 at 4. ALJs we spoke with doubted this confidence in technology. *E.g.*, ALJ 2 at 4; ALJ 21 at 7. No ALJ, for instance, agreed with the confidence SSA has placed in the eBB as a tool to increase quality decision making without a sacrifice in quantity. Management literature on goal setting posits an irreducible tension between quality and quantity when tasks involve complex decision-making. *E.g.*, Stephen W. Gilliland & Ronald S. Landis, *Quality and Quantity Goals in a Complex Decision Task: Strategies and Outcomes*, 77 J. APPLIED PSYCH. 672, 680 (1992). The agency’s Inspector General recently reported a dropoff in ALJ productivity and explained this decline as produced in part on renewed efforts at quality control. OIG, EFFORTS TO ELIMINATE THE HEARING BACKLOG, *supra* note 69 at 3-4.

⁴⁰² *Cf. Nash v. Bowen*, 869 F.2d 675, 680 (2d Cir. 1989).

⁴⁰³ *E.g.*, David A. Fahrenthold, *At Social Security Office With a Million Person Backlog, There’s a New Chief*, WASH. POST, July 23, 2015.

agency has the complex task of managing a complex balance between quantity and quality as successfully as possible.

The federal courts do not shoulder anywhere near the same obligation to generate decisions quickly,⁴⁰⁴ although some districts have quite sizeable dockets. Moreover, the federal courts do not endure the same legislative and public scrutiny for their pace of decision-making. Federal judges can render particularized justice tailored to the circumstances of an individual case, without significant regard for production quotas.

To a certain extent different resources allotted to each institution determine these contrasting goals. While federal courts cannot spend limitless time deciding any particular motion, many judges and their clerks take longer on a social security claim than what all personnel at the ALJ and Appeals Council levels can spend in total. One federal judge we asked estimated that she spends two to eight hours on a social security appeal.⁴⁰⁵ But all others reported that they spent longer, ranging from 15-25 hours to 60-80 hours of judge and clerk time.⁴⁰⁶ In contrast, the huge numbers of claims permit ALJs and decision writers to combine to average about ten or eleven hours on a claim, if our interviews are any guide; a claim may get five hours of adjudicator attention at the Appeals Council.⁴⁰⁷ “We have to rob Peter to pay Paul,” another ALJ told us, while district judges “speak from a place of unlimited

⁴⁰⁴ Cf. POSNER, *supra* note 394 at 140-141 (commenting on district judges’ sensitivity to delays in deciding motions but noting that there is no sanction for delays). When we described ALJ caseloads to one district judge, her response was “that’s preposterous.” Federal Judge 4 at 5.

⁴⁰⁵ Federal Judge 1 at 3. The complaint filed in 2013 in a case challenging the ALJ productivity goal described federal judges as spending 4.6 hours to adjudicate an agency appeal. Complaint, Association of Administrative Law Judges v. Colvin, Civ. No. 13-2925, N.D. Ill., Apr. 18, 2013, at 19 (hours per claim). This figure is inconsistent with what all but one federal judge reported to us.

⁴⁰⁶ Federal Judge 19 at 4 (15-25 hours); Federal Judge 4 at 4 (60-80 hours); Federal Judge 10 at 3 (30-40 hours); Federal Judge 15 at 4 (15-20 hours); Federal Judge 11 at 4 (two days for law clerk, ten days for judge); Federal Judge 21 at 3 (2-3 hours to review law clerk’s draft, several days for clerk to draft opinion); Federal Judge 20 (1-2 weeks); Federal Judge 6 at 1 (at least 30 hours); Federal Judge 7 (one week of judge and clerk time).

⁴⁰⁷ As discussed in Part II, ALJs reported to us that they spent about two-and-a-half hours on a claim. Decision writers spend about eight hours on a decision denying benefits. An analyst at the Appeals Council is supposed to prepare two cases per day, and an AO and AAJ to work through 5-12 cases per day.

possibilities.”⁴⁰⁸ Also, federal judges make extensive use of law clerk resources in social security cases, with many judges assigning these appeals to permanent clerks. These staff handle a large number of cases and thereby develop deep familiarity with the applicable law and medical records, narrowing the expertise gap between the federal judiciary and ALJs. In contrast, the decision writer-to-ALJ ratio is often 1:1, and ALJs for the most part do not have the same individualized relationships with decision writers that federal judges enjoy with their clerks.⁴⁰⁹

An ALJ adjudicating 40-50 cases per month and a district judge and her clerk “picking apart a case for a week” have fundamentally different jobs, as an ALJ aptly put it.⁴¹⁰ ALJs spend much more time with medical records and the law of social security benefits than federal judges do, and for this reason are surely more expert in the handling of medical evidence and the legal regime’s application. Given the governing law’s endless details, the lengthy sets of medical records claimants often submit, and ALJs’ duty to develop the record, however, a federal judge spending three times the time an ALJ has for a decision can almost invariably find deficiencies. The amount of time different types of decision makers have to deliberate is an institutional variable that has nothing to do with the quality of decision maker.⁴¹¹

2. Different Legal Commitments

The roles that ALJs and federal judges play contrast in another important respect. This difference also produces some portion of remands without either institution malfunctioning. The SSA has to administer a national program, one for which consistency is an obviously desirable

⁴⁰⁸ ALJ 1 at 5; *see also* ALJ 18 at 4 (district judges are “in a different world”); ALJ 20 at 3 (suggesting that ALJs are “a little bitter” about the difference in time between district courts and hearing offices); ALJ 3 at 4 (commenting that district judges have good knowledge of applicable law but no understanding of the “problems” ALJs face when handling a case).

⁴⁰⁹ Staff-to-ALJ ratios bear importantly on ALJ productivity. OIG, ADMINISTRATIVE LAW JUDGE AND HEARING OFFICE PERFORMANCE, *supra* note 96 at 5.

⁴¹⁰ ALJ 4 at 2.

⁴¹¹ Henry J. Friendly, *Indiscretion About Discretion*, 31 EMORY L.J. 747, 757 (1982).

goal.⁴¹² To say nothing of legitimacy concerns, the agency would weather well-deserved political blowback if it treated claimants in Arizona more favorably than claimants in New Mexico because of differences between the Ninth and Tenth Circuits. Moreover, careful attention to the nuances of circuit or district court case law within the agency might create significant administrative headaches. National case assistance centers, National Hearing Centers, and remote adjudication by video are possible in part because the SSA counsels ALJs and decision writers against citing case law in decisions.

District and magistrate judges, in contrast, must follow their circuit's precedent. The task of harmonization belongs to the circuits and the Supreme Court. Inconsistency in the application of federal law is routine in the ninety-four federal districts. Courts of appeals have more of an obligation to inter-circuit harmony when they develop a federal legal regime in their decisions.⁴¹³ But the degree to which this obligation trumps others is a contested jurisprudential question with no clear answer.⁴¹⁴ Even if no applicable precedent exists, a district court is unlikely to treat case law from within the same district cavalierly and rather will accord it persuasive force.⁴¹⁵

These different legal commitments manifest themselves in a number of ways that necessarily generate remands. For instance, when OGC lawyers ask the Appeals Council to agree to a voluntary remand, they must explain their reasons in terms of the statute, regulations, and Social Security rulings, not case law.⁴¹⁶ For the most part, neither OGC lawyers nor Appeals Council personnel we interviewed told us that they consider either the district where the

⁴¹² *E.g.*, HUME, *supra* note 273 at 99.

⁴¹³ *E.g.*, Renteria-Gonzalez v. INS, 322 F.3d 804, 814 (5th Cir. 2002).

⁴¹⁴ See generally Martha Dragich, *Uniformity, Inferiority, and the Law of the Circuit Doctrine*, 56 LOY. L. REV. 535 (2010).

⁴¹⁵ *E.g.*, Warneka v. Colvin, Civ. No. 14-22, 2015 WL 1470955, at *2 n.4 (D.N.H. Mar. 31, 2015) (rejecting SSA's invitation to disregard a district court opinion as inconsistent with the agency's interpretation of an SSR); Rossiter v. Astrue, Civ. No. 10-349, 2011 WL 2783997, at *2 (D.N.H. July 15, 2011).

⁴¹⁶ Agency Official 6, First Interview at 2.

case proceeds or the judge to whom it is assigned when they make RVR determinations.⁴¹⁷ An approach to RVRs more finely tailored to the idiosyncrasies of a particular court may reduce remands. Some circuits hammered the agency over boilerplate that ALJs routinely included when evaluating the credibility of a claimant’s alleged symptoms.⁴¹⁸ The Eighth Circuit was more forgiving.⁴¹⁹ Were case law the proper determinant, an OGC lawyer in the Northern District of Illinois could have had a different RVR threshold than an OGC lawyer in the Eastern District of Arkansas. But these different thresholds would have undermined the agency’s commitment to national uniformity.

The agency has instructed ALJs and decision writers “not to consider any district court decisions” as sources of legal guidance when making their decisions.⁴²⁰ This admonition is consistent with the agency’s longstanding policy not to treat appellate decisions as binding unless the SSA issues an acquiescence ruling.⁴²¹ The reasons for this policy are several and include both the agency’s commitment to a national program and the administrative complexity of accounting for district court nuances when drafting and revising decisions.⁴²² However justified, the policy surely generates remands. A district or magistrate judge will likely afford

⁴¹⁷ Agency Official 6, First Interview at 2; OGC Lawyer 12 at 2; OGC Lawyer 8 at 4; OGC Lawyer 24 at 3. One agency official explained that, in order to grant an RVR request, the agency needs to be able to translate critiques expressed in terms of case law into deficiencies rooted in the statute and regulations. Agency Official 6, Second Interview at 4.

⁴¹⁸ The boilerplate was the following:

After careful consideration of the evidence, the undersigned finds that the claimant’s medically determinable impairments could reasonably be expected to cause the alleged symptoms; however, the claimant’s statements concerning the intensity, persistence and limiting effects of these symptoms are not credible to the extent they are inconsistent with the above residual functional capacity assessment.

For this language, and for an example of an opinion criticizing it, see, e.g., *Bjornson v. Astrue*, 671 F.3d 640, 644-646 (7th Cir. 2012). The agency changed this text in 2012 due to this sort of criticism.

⁴¹⁹ *Kamann v. Colvin*, 721 F.3d 945, 951-52 (8th Cir. 2013).

⁴²⁰ Memorandum to All Administrative Law Judges and All Senior Attorneys from Debra Bice, Chief Administrative Law Judge, Jan. 11, 2013, at 2 (on file with authors).

⁴²¹ Soc. Sec. Ruling 96-1p.

⁴²² Bice Memorandum, *supra* note 420 at 2; Decision Writer 3 at 4; ALJ 17 at 3.

decisions rendered by colleagues persuasive force.⁴²³ Just by mentioning certain well-known cases, some ALJs believe, they could insulate their decisions from a court-initiated remand.⁴²⁴

3. Different Perspectives

A final contrast that generates claimant wins has to do with perspectives that differ along two institutional axes. The first has to do with baselines. ALJs handle a much larger sample of cases each year than federal judges, and ALJs get their cases earlier in the adjudication process. Presumably the ALJ sees a wider array of types of impairments, and the ALJ gets many more slam-dunk cases. The easy ones should get weeded out well before they reach district court. An ALJ may therefore have a different “cutpoint”⁴²⁵ – roughly, the line the ALJ would draw along a given dimension between disability and no disability – than a federal judge for what she believes qualifies as a disability.⁴²⁶ These different cutpoints, produced by different baselines, will give ALJs and federal judges institutionally determined understandings of disability that differ. One ALJ who had previously served as an OGC lawyer described this phenomenon aptly to us. What seemed like a “slam dunk” claimant win at OGC changed when she became an ALJ. “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.”⁴²⁷

Another baseline difficulty involves the evaluation of expert and lay evidence. With their immense case loads, ALJs and decision writers told us that they routinely see letters from the

⁴²³ *E.g.*, *Warneka v. Colvin*, Civ. No. 14-22, 2015 WL 1470955, at *2 n.4 (D.N.H. Mar. 31, 2015) (rejecting SSA’s invitation to disregard a district court opinion as inconsistent with the agency’s interpretation of an SSR); *Rossiter v. Astrue*, Civ. No. 10-349, 2011 WL 2783997, at *2 (D.N.H. July 15, 2011).

⁴²⁴ *E.g.*, ALJ 10 at 2. An OGC lawyer insisted to us that the agency is “going to get cases sent back” because of its exclusionary rule for district court case law. OGC Lawyer 10 at 3.

⁴²⁵ *Masur*, *supra* note 392 at 483.

⁴²⁶ We appreciate that most ALJs and federal judges make their decisions based on a good faith effort to apply the law to their best understandings of the facts. Within the legal regime for disability determinations, however, there is room for judgment calls. Also, we expect that judges’ impressions of claimants and their needs influence decision-making, however consciously. One federal judge we interviewed candidly admitted that this was so. Federal Judge 24 at 1. For these reasons, where an adjudicator’s cutpoint lies is important.

⁴²⁷ ALJ 20 at 1. For a discussion of this issue, see *MASHAW ET AL.*, *supra* note 28 at 138-139.

same physicians that use the same phrases to describe patients with strikingly similar problems.⁴²⁸ “We know which doctors are trustworthy and which ones aren’t,” one ALJ told us, “but we can’t put this in a decision.”⁴²⁹ But presumably ALJs will discount these physicians’ opinions in a policy-compliant manner.⁴³⁰ Likewise, one ALJ told us, claimants can testify in an obviously coached manner, taught to say just the right thing to buttress a claim for benefits.⁴³¹ Federal judges have nowhere near the sample of cases to draw upon and thus may have less capacity to identify suspect medical evidence or embellished claimant testimony. The ALJ’s baseline might breed undue skepticism, a concern raised by a number of ALJ comments we heard.⁴³² The federal judge’s baseline, in contrast, might foster unwarranted credulity.

Institutional perspectives clash along a second axis. ALJs are supposed to conduct hearings in an inquisitorial manner.⁴³³ As federal judges understand this duty, it finds expression in an obligation to develop the record,⁴³⁴ or otherwise to explore the evidence to the claimant’s benefit.⁴³⁵ To ALJs, however, the inquisitorial duty has a different salience. Generally speaking, no attorney represents the government in benefits hearings, while more than three-fourths of claimants have representation. ALJs must aid claimants with their presentations, to be sure. But without a government representative, an ALJ must also probe for inconsistencies and otherwise test the adequacy of a claim for benefits. One ALJ described the task as to “protect the

⁴²⁸ ALJ 1 at 5; ALJ 21 at 4.

⁴²⁹ ALJ 7 at 4; Decision Writer 14 at 6; ALJ 6 at 4. *Cf.* *Lester v. Chater*, 81 F.3d 821, 832 (9th Cir. 1995) (holding that an ALJ may not assume that a physician is lying to help his or her patient).

⁴³⁰ Another ALJ described claimant firms sending claimants to the same doctors and getting the same statements from treating sources. ALJ 16 at 3.

⁴³¹ ALJ 1 at 5 (reporting that claimants repeatedly testify that they can only lift “about a gallon of milk”); ALJ 6 at 4 (describing how claimants know to say that they can only lift “a gallon of milk”).

⁴³² The ALJ who discussed this issue at length with us insisted that “people are trying to scam all the time.” ALJ 1 at 6.

⁴³³ *See generally* *Sims v. Apfel*, 530 U.S. 103, 110-111 (2000).

⁴³⁴ *E.g.*, *Butts v. Barnhart*, 388 F.3d 377, 386 (2d Cir. 2004); *Blea v. Barnhart*, 466 F.3d 903, 911 (10th Cir. 2006).

⁴³⁵ *E.g.*, *Overman v. Astrue*, 546 F.3d 456, 465 (7th Cir. 2008); *Steele v. Barnhart*, 290 F.3d 936, 941 (7th Cir. 2002).

fund.”⁴³⁶ A federal judge, presiding over an adversarial process, can rely on the OGC lawyer to make the case that a claimant is malingering or otherwise not entitled to benefits. ALJs shoulder this duty themselves, one merged uncomfortably with their duty to aid the claimant with their case and serve as an objective fact-finder.⁴³⁷ Several ALJs described to us an ingrained skepticism to claimants,⁴³⁸ an attitude inconsistent with agency policy but one understandable in light of the several hats ALJs must wear. With no obligation to ferret out defects in a plaintiff’s case, a federal judge has no institutional reason to approach cases with such a bent.

* * *

It is possible that the factors that contribute to the remand rate have more to do with conflicts between two different institutions than anything else. Unless one could say with confidence that one institution’s commitments, goals, and perspectives are normatively inferior to the other’s, judgments about the remand rate, viewed statically, are hard to draw.

Only Congress can narrow significantly the institutional gaps between the agency and the courts. Funding permitting a dramatic increase in the ALJ corps, for example, would lessen the tension between quality and quantity by giving ALJs the sort of time that federal judges enjoy to decide claims. It would be misguided to read into the overall remand rate something normatively significant about how the agency adjudicates claims, or about the standard by which federal courts measure ALJ decisions, without appreciating the demands Congress makes of those who administer the social security disability program.

⁴³⁶ ALJ 15 at 1. We do not understand this ALJ as expressing some felt obligation to deny claims in order to maintain fiscal rectitude. Rather, his statement reflects an awareness that the agency as a whole, and he as its representative, have an obligation to be a responsible steward of social security funds. Cf. Lael R. Keiser, *Understanding Street-Level Bureaucrats’ Decision Making: Determining Eligibility in the Social Security Disability Program*, PUB. ADMIN. REV., Mar./April 2010, at 247, 250 (“DDS offices have two contradictory missions. The first is to grant access to a needy population, and the second is to protect the solvency of the Social Security Trust Fund.”).

⁴³⁷ ALJ 19 at 2. To one ALJ we interviewed, it is “totally unfair” that he has both to cross-examine the claimant and to maintain a non-adversarial atmosphere. ALJ 19 at 2. See also Bloch et al., *supra* note 94 at 7.

⁴³⁸ ALJ 8 at 3; ALJ 14 at 1; ALJ 19 at 2.

Part IV. Variations Across Districts

A snapshot of the national remand rate means little. Perhaps changes in the remand rate over time might reveal more about how well or poorly the agency or the federal courts are functioning, at least if the composition of claims remained unchanged. But the large swings in the labor market that began with the Great Recession make us doubt this story of consistency over time. Thus, we believe trends in the remand rate are of limited use in understanding deeper facts about the disability appeals process.

With this proviso, we note that while the remand rate dropped steadily from above 60% in 2001 to 45% in 2014, it has done so steadily and relatively slowly, rather than abruptly. That fact suggests that the trend in the national remand rate is unlikely to be the result of cataclysmic economic forces, most notably including a weakening of claims due to the Great Recession's onset. It is likely true that as the labor market worsens, more people choosing between employment and disability benefits might pursue the latter instead of the former. But there are at least two problems with the supposition that the national trend in the remand rate is simply the result of national macroeconomic forces. First, one would expect a major drop in the remand rate in 2011 or 2012, when claims initially filed in 2007-2008 hit the district courts; no such major drop actually occurred.⁴³⁹ Second, one would also expect to see the remand rate ticking up thereafter, corresponding to the period roughly four years after the labor market improved following the worst months of the recession; no such change has so far shown up in the data.

A more intriguing story couples changes in claim composition with structural changes within the agency that have improved the overall quality of claims that federal courts review. ODAR began to implement quality assurance initiatives in earnest around 2011, roughly around the time we would expect the average quality of cases appealed to the district courts to begin to

⁴³⁹ See discussion in note 280 at 45, *supra*.

rebound. Such a change in case quality might have counteracted any effects of the agency's quality assurance initiatives. The result would be a stronger set of claims facing a better-performing agency review process, such that recent improvements in the agency's multi-layered review would dampen whatever case selection effects are caused by labor market and demographic swings.

A third story is related but different. Perhaps the multi-layered review within the agency ensures that the composition of claims that finds their way into the district courts would be relatively unaffected by broader economic and social forces. Put metaphorically, the mesh in the Appeals Council's net is fine enough to catch most erroneous ALJ decisions, regardless of initial claim quality. On this story, the modest decline in the federal courts over the past decade does not reflect the influence of exogenous variables related to case quality, but rather factors endogenous to the federal judiciary, such as changes in judicial personnel, changes in some key circuit doctrine, and so forth.

Our data allow us to comment on the first account, but they are insufficient to enable us to investigate the second or third story in detail. To do so, we would need, at a minimum, detailed data concerning initial claims and appeals at every level of the agency review process, over a period of years. We lack such data. This fact necessarily limits the conclusions we do reach, as discussed below.

In this Part we focus on the more modest objective of characterizing and trying to understand the correlates of district-level variation in remand rates. The institutional factors we discussed in Part III that might be expected to produce an irreducible core of claimant wins – different legal commitments, different goals, and different perspectives – do not vary from one

district to the next. Yet outcomes in social security appeals differ strikingly across the 94 judicial districts.

The following table reports the district-specific remand rate among all cases in our OGC data for the years 2010-2013.⁴⁴⁰ The table shows substantial variation across districts—from a low of 20.8% in the Eastern District of Arkansas to a high of 76.0% in the Southern District of New York. The median value was 41.4%, which occurred in the Middle District of Florida. Formal statistical tests easily reject the null hypothesis that the district-level remand rate is the same across districts, meaning that there is convincing statistical evidence that district-level variation is not simply the product of random variation.⁴⁴¹

⁴⁴⁰ As we discuss in the Data Appendix, the data we refer to as our “OGC data” were provided by the agency for purposes of this report and originate from the NDMIS system maintained by OGC. The calculations discussed in the report were made on a substantial subset of these data, whose construction we describe in the Data Appendix; see page A-3 of that appendix for discussion of why we did not generally use alternative data from ODAR that the agency provided. While we are aware of no reason to believe this subset is unrepresentative of the overall set of data, we emphasize that official agency statistics and those we created using our OGC data should not be expected to overlap perfectly.

⁴⁴¹ A simple way to measure the importance of district-level variation is to estimate a least-squares model in which the outcome variable is a dummy variable indicating whether a case was remanded, and the dependent variables are a set of dummy variables for the district courts, with each dummy variable indicating whether a case was filed in the corresponding district. The estimated constant will equal the remand rate for the excluded district, with the coefficients on the dummy variables equaling the deviation from the excluded district’s mean for the district corresponding to each dummy variable. A test of whether the remand rate varies systematically across districts is then equivalent to testing the null hypothesis that all the district dummy variables’ coefficients equal zero, which can be implemented via a standard asymptotic chi-square test. We found that that easily rejects the null hypothesis, which means that there is statistically significant evidence that the underlying remand rate varies across districts.

Table 1: District-Level Remand Rates for Cases Decided 2010-2013⁴⁴²

State District		Remand Percentage	State District		Remand Percentage	State District		Remand Percentage
AR	E	20.8	NC	W	36.5	RI		49.2
WV	S	21.2	TN	E	37.0	AR	W	49.4
MS	S	23.0	MD		37.1	CA	C	50.6
KY	E	23.9	TN	M	37.2	OK	W	52.8
TX	W	24.0	MI	E	37.5	NY	N	53.1
MT		24.5	CA	E	37.6	KS		54.0
FL	N	24.6	MO	W	38.5	NJ		54.1
AL	N	26.7	TX	E	38.6	SC		54.2
VA	E	27.2	MI	W	38.6	SD		54.5
KY	W	27.4	PA	M	38.9	IN	N	54.9
AL	M	27.5	DC		39.1	IA	N	55.1
NV		27.6	CA	N	39.9	OR		55.4
VA	W	27.8	TX	N	39.9	GA	N	55.4
TX	S	27.8	ME		39.9	CO		56.4
ND		30.4	FL	M	41.4	NY	W	56.9
GA	M	30.7	WY		42.2	DE		57.1
IA	S	31.0	IL	C	42.4	NC	E	58.4
WV	N	31.4	WA	E	42.6	PA	E	59.2
MN		32.7	MS	N	42.9	NH		61.1
TN	W	32.9	LA	W	42.9	IL	S	63.1
HI		33.3	PR		44.2	CT		63.9
FL	S	33.4	MA		44.5	VT		65.4
PA	W	33.6	IN	S	45.5	AK		65.5
OH	N	33.9	OH	S	46.6	IL	N	67.4
CA	S	33.9	NC	M	47.7	NM		69.2
GA	S	35.4	ID		48.5	WA	W	69.2
NE		35.6	AZ		48.8	NY	E	71.9
MO	E	36.2	OK	N	49.1	OK	E	75.9
						NY	S	76.0

Outcome consistency, or the notion that like cases ought to be treated alike, is an important value that any system of adjudication ought to serve.⁴⁴³ ALJs have earned criticism

⁴⁴² The data we used for this table permit us to report remand rates for only eighty-five of the ninety-four districts, for reasons we explain in the Data Appendix.

⁴⁴³ Yoav Dotan, *Making Consistency Consistent*, 57 ADMIN. L. REV. 995, 1000-1001 (2005) (explaining the value of consistency in administrative decision-making); Stephen H. Legomsky, *Learning to Live With Unequal Justice: Asylum and the Limits to Consistency*, 60 STAN. L. REV. 413, 423-428 (2007).

for past failures to honor this value.⁴⁴⁴ Do federal judges deserve similar criticism? Unlike ALJs, federal judges do not owe fidelity to a single body of nationally uniform law. But, while circuit boundaries explain a lot of district-level variation, they leave some differences unexplained. Why do claimants in Albany (53.1% remand rate) win much less often than claimants in Brooklyn (71.9% remand rate)? Should Pasadena claimants (50.6% remand rate) prevail more often than their counterparts in Fresno (37.6% remand rate)? These questions helped motivate our attempts to understand the variation in district-level remand rates.⁴⁴⁵

Section A provides quantitative evidence concerning the nature of district-level variation, with a focus on attempting to determine what factors are associated with this variation. We first determine that district-level variation in remand rates is not driven by a small number of outlier districts.⁴⁴⁶ We then consider two hypotheses we developed based on our qualitative investigation. First, our interview subjects routinely chalked up differences in outcomes to individual judicial preferences. Does the behavior of “single outlier judges”—those who are very influential in certain districts—explain why some districts have especially high or low remand rates by comparison to the national average? Our evidence indicates that judges tend to march in lockstep with other judges in the same district, not to the beat of different drummers: while considerable variation across judges does exist, the single outlier judge hypothesis does little to explain districts’ remand rates overall.

⁴⁴⁴ HAROLD J. KRENT & SCOTT MORRIS, ACHIEVING GREATER CONSISTENCY IN SOCIAL SECURITY DISABILITY ADJUDICATION: AN EMPIRICAL STUDY AND SUGGESTED REFORMS, April 2013, at 1. So have immigration judges. Jaya Ramji-Nogales et al., *Refugee Roulette: Disparities in Asylum Adjudication*, 60 STAN. L. REV. 295, 296 (2007).

⁴⁴⁵ The agency has noted to us that DDS-level allowance rates differ across areas, and also that available medical treatment might vary. However, if the agency’s multi-layered review process operates in a uniform way, then one would expect such variation to have been appropriately accounted for in that process: claims that a state’s DDS process incorrectly denies should be allowed by the agency itself on review. In any case, the agency did not provide us data from the DDS level forward, so this is not an issue we can address quantitatively in this report.

⁴⁴⁶ Limitations in the OGC data prevent us from identifying cases in nine districts, including the Eastern District of Wisconsin, whose remand rate is high compared to the national average. However, our data represent an overwhelming majority of district court cases, so including these districts, if we could, would not likely affect our basic conclusion on this point.

The second hypothesis is that district-level variation is explained by circuit membership. Harold Krent and Scott Morris reached this conclusion in their study of district court remand variations, concluding that “the doctrine or ‘culture’ within a particular judicial circuit makes a substantial difference”⁴⁴⁷ Our results comport with theirs and reveal substantial evidence of a circuit-level component to district-level variation in remand rates. But we also find that considerable variation in district-level remand rates exists even *within* circuits.

District-level factors that are both broader than individual judges and narrower than circuit membership must play an important role in explaining the variation in remand rates across districts. This fact is the key puzzle that needs explaining. Based on an intensive qualitative study of all aspects of the social security claims adjudication process, with a particular focus on three districts, we identified a set of candidates that might answer this puzzle: caseloads, judicial ideology, and labor market variations. Very few of these are significant, as we report in Section B. We did, however, discover a couple of interesting phenomena that our other factors do not explain. To the extent we can measure with the data used here, decision-making within the agency seems to differ geographically. We combine these findings with ones from our qualitative investigation of three districts in Sections C and D to hypothesize, albeit tentatively, that district culture, determined at least in part by the quality of ALJ decisions federal judges review, likely explains a component of district-level variation. Section E summarizes our findings as to variation across districts in the remand rate.

A. A Quantitative Analysis of District-Level Variation in Remands

In this section we attempt to disentangle judge-, district-, and circuit-level explanations for the variation in district-level remand rates. We note at the outset that the lines among these types of explanations may blur. An example much on the mind of OGC lawyers we interviewed

⁴⁴⁷ Krent & Morris, *supra* note 1 at 367.

offers a useful illustration. Compared to the national average, the Eastern District of Wisconsin has a high remand rate. One judge in particular has criticized ALJ decisions.⁴⁴⁸ To a certain extent, his frustration reflects inconsistencies between ALJ decisions and case law from the Seventh Circuit,⁴⁴⁹ a comparatively claimant-friendly court. But, an OGC lawyer explained to us, ALJ decisions of questionable quality – a fact of life for all judges in this district, if the OGC lawyer is right – have also fueled this judge’s reactions.⁴⁵⁰ According to that OGC lawyer, then, this district’s remand rate is a function of an individual judge’s preferences, prevailing law and tendencies within the circuit, and the quality of inputs for district court decision-making.⁴⁵¹

1. The Nature of the Problem of District-Level Variations

An obvious and potentially important difference across districts involves differences in the nature of claimants’ alleged disabilities. For example, the Southern District of New York is unlikely to have many coal miners claiming benefits, just as districts located in Appalachia are unlikely to have many cab drivers. More rural districts will have fewer manufacturing workers claiming, while heavily urbanized districts will have fewer farmers filing for disability benefits. It is possible that these or other area-related sources of heterogeneity in claimant characteristics could lead to differences in the accuracy of the agency’s adjudication of claims. If so, then district-level variation in claim quality might be a source of district-level variation in the remand rate. We are unable to control for such claimant-level sources of heterogeneity. We assume for purposes of this report that this inability does not render the rest of our quantitative analysis

⁴⁴⁸ *E.g.*, Jane Pribek, *Federal Judges Fired Up Over Social Security Cases*, Wis. L.J., Mar. 11, 2013.

⁴⁴⁹ In a scheduling hearing held on March 13, 2013, the district judge complained of “far too many administrative law judges who are not conversant in Seventh Circuit law.” Transcript, Mar. 13, 2013, at 14, *at* www.ssaconnect.com/260-sanctions.

⁴⁵⁰ OGC Lawyer 1 at 5 (complaining of “ill-trained, slapdash ALJs” generating decisions for this judge to review).

⁴⁵¹ We note again that the Eastern District of Wisconsin could not be identified in the OGC data we use for much of this Part. We use it anecdotally in the text above to illustrate a general point.

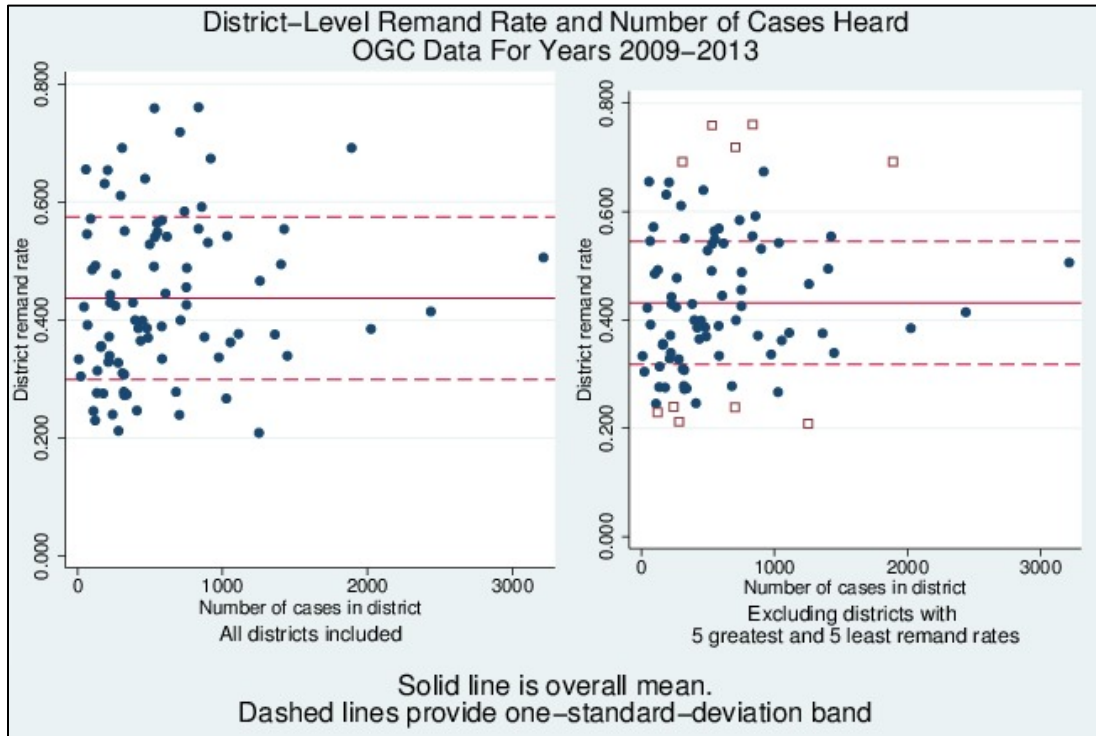
unuseful. Still, we urge readers of the quantitative sections of this report to keep that limitation in mind.

With that caveat noted, variation in district-level remand rates would have much less systemic importance if it were largely the product of a few outlier districts, with most districts having reasonably similar rates. One way to test this proposition is to determine what happens when one eliminates outlier districts, or those with remand rates that deviate greatly from the average in either direction. Also, variations are less systemically concerning if small districts with very few cases are the outliers, because smaller districts will tend to have more variable remand rates, other things equal. To assess whether variation in district-level remand rates is primarily driven by the fact that some districts have fewer cases than others, we plotted each district's remand rate over the 2010-2013 period against the number of cases from which this rate was calculated.

The resulting graph appears in the left panel of Figure 1 below. The district-level mean of 43.7% is shown as the solid line toward the middle of the graph. The two dashed lines above and below this line indicate the values of the remand rate that are one standard deviation above and below the mean. These are useful because a unit of standard deviation is a common measure of a sizable, but not extreme, movement in a variable's value.⁴⁵² The standard deviation for all districts was 13.8%, so the one-standard deviation band consists of districts with remand rates between 29.9% and 57.5%. The distribution of districts outside the one-standard deviation band is not precisely symmetric, as remand rates among high-remand districts are farther from the mean than for those in districts with low remand rates.

⁴⁵² The standard deviation of a statistic is the square-root of its variance.

Figure 1



The panel on the right side of Figure 1 includes the same plot, with two adjustments. First, hollow squares represent the remand rates of those districts whose remand rates are among either the five highest or five lowest.⁴⁵³ Second, the middle horizontal line and the associated one-standard deviation band in the panel on the right side of Figure 1 are calculated excluding the top-5 or bottom-5 remand rate districts.⁴⁵⁴ The mean across districts in this panel was 43.2%, virtually identical to the mean of 43.7% for the left panel. The distance of the remand rates

⁴⁵³ The districts with the five lowest remand rates are the Eastern District of Arkansas, with a remand rate of 20.8% (1252 cases); the Southern District of West Virginia, 21.2% (283 cases); the Southern District of Mississippi, 23.0% (122 cases); the Eastern District of Kentucky, 23.9% (703 cases); and the Western District of Texas, 24.0% (242 cases). The districts with the five highest remand rates are the Southern District of New York, with a remand rate of 76.0% (835 cases); the Eastern District of Oklahoma, 75.9% (531 cases); the Eastern District of New York, 71.9% (707 cases); the Western District of Washington, 69.2% (1891 cases); and the District of New Mexico, 69.2% (308 cases).

⁴⁵⁴ The one-standard deviation band for district-level remand rates runs between 31.8% and 54.5% for the right panel. Given the virtually identical means in the two panels, this result is an arithmetic necessity, because eliminating outlying points necessarily eliminates those points that contribute the greatest values to the summand in computing the standard deviation. Thus the simple fact of a tightening itself carries no important information; to be meaningful, the drop in the standard deviation following elimination of the most extreme outlier points would have to be very substantial. The actual standard deviation in the right panel is 11.5%, which we do not think is substantially different from the left pane's 13.9%.

among the top-5 and bottom-5 districts from the overall mean thus is not so great as to meaningfully affect the overall mean.⁴⁵⁵ Variance in districts' remand rates, in short, does not chiefly result from a few outliers.

Figure 1 yields a final insight. Eight of the districts with the ten largest numbers of cases in our data are within the left panel's one-standard deviation band.⁴⁵⁶ These districts accounted for more than a quarter of the cases represented in our data for the 2010-2013 period. Huge caseloads do not seem to push districts to either extreme.

An inquiry into what causes district-level variations would yield considerably different conclusions if districts' remand rates changed dramatically from year to year. To address this possibility, we estimated a regression model in which the dependent variable was each district's remand rate for a given year. The lone regressor was the value of the district's remand rate in the preceding year, or what is known in statistical parlance as the "first lag" of the dependent variable.⁴⁵⁷ If the remand rate is persistently high in some districts and persistently low in others, then we would expect a coefficient estimate for the preceding year's remand rate to be substantially above zero. In the extreme situation in which each district always has the same remand rate in every year, the coefficient for the preceding year's remand rate would always be exactly 1.

⁴⁵⁵ For top-5 districts, the remand rate was 72.4%, roughly 29 percentage points above the overall mean of 43.7%. For the bottom-5 districts, the remand rate was 22.6%, about 21 percentage points below the overall mean.

⁴⁵⁶ The top ten districts in terms of number of decided cases were the Eastern District of Arkansas (1252 cases); the Southern District of Ohio (1259 cases); the Eastern District of Michigan (1362 cases); the Western District of Arkansas (1404 cases); the District of Oregon (1426 cases); the Northern District of Ohio (1448 cases); the Western District of Washington (1891 cases); the Western District of Missouri (2025 cases); the Middle District of Florida (2438 cases); and the Central District of California (3215 cases). Of these, only the Western District of Washington and the Eastern District of Arkansas had remand rates—69.2% and 20.8%, respectively—that were outside the one-standard deviation band.

⁴⁵⁷ This regression model has 255 observations—85 districts for which we had useable data, times 3 annual observations per district (since we are using only data from 2010-2013, data for 2010 enter only as first lags).

This model generated a coefficient on the preceding year's remand rate of 0.77.⁴⁵⁸ On average, districts with a remand rate that is greater by one percentage point in a given year have a remand rate that is greater by roughly three-quarters of a percentage point in the following year. We emphasize that this is not a causal claim: we are not suggesting that any factor that increases a district's remand rate by one point in one year will necessarily increase the following year's remand rate by 0.77 points. Rather, the described relationship is merely associational. Nevertheless, it does tell us that district-level remand rates are persistent over time.

In sum, the results of this section have shown that there is a substantial degree of year-to-year persistence over time in district-level remand rates. Further, while district-level remand rates vary systematically, the overall remand rate is not driven by a small number of outlier districts.

2. Are Outlier Judges Important?

Differences in district-level remand rates have a different institutional significance if they result from individual judge idiosyncrasy. The policy implications of a situation where outlier judges push their districts' remand rates away from national- or circuit-average levels differ considerably from a situation where rates are less sensitive to individual judicial preference.

Idiosyncrasies in the remand decisions of individual judges might be expected to produce dramatic variations in remand rates when districts are small or when they decide few social security cases. The impact of a single outlier judge on a small district's remand rate might be enough to push the district's remand rate far away from the average remand rate for the circuit in which the district is located.

⁴⁵⁸ This estimate is precisely enough estimated that we can rule out the possibility that the coefficient is either zero or one; thus there is substantial persistence, but a bit less than would be expected if nothing changed from year to year at the district level.

A district that came up repeatedly in our interviews (“District A”) offers an example.⁴⁵⁹ District A had a remand rate about 20 percentage points higher than the average remand rate of all districts within its circuit. One of the judges within District A remanded cases a rate 10 percentage points higher than District A’s average. This judge accounted for more than a third of all decisions rendered by District A in our data. If this outlier judge’s decisions are removed from consideration, however, District A’s remand rate falls by only 6.7 percentage points. District A’s rate without the outlier judge included, in other words, is considerably closer to District A’s rate with the outlier judge included than it is to the circuit’s rate without District A included. Thus even a judge with a substantial share of the district’s cases, and with a reputation as an outlier, might account for relatively little of the value of his district’s remand rate.

District A is not unique in this respect. To determine the impact of individual judges on district remand rates, we identified the judge assigned to cases using the “judge” field provided in dockets.⁴⁶⁰ Unless specifically indicated, for purposes of this analysis we refer to magistrates and Article III judges collectively as “judges.”⁴⁶¹ Using this definition, we found a total of 1,369 judges in our data.⁴⁶² Across all of these judges, the average remand rate was 45.8% (the median was similar, at 44.4%). The remand rate varied widely. For example, 101 matched judges had a remand rate of zero and 88 had a remand rate of 100%. Not surprisingly, these extremes were associated with judges who decided relatively few cases—an average of 2.3 and 1.9 cases in each of these two categories, respectively. Even among those with remand rates between zero

⁴⁵⁹ We have kept the district anonymous at the agency’s request.

⁴⁶⁰ See the Data Appendix for more detail on the construction of our matched OGC-dockets data.

⁴⁶¹ In some cases there is a name in the “case referred to” field as well; typically this name appears to be that of a magistrate judge.

⁴⁶² As discussed in the Data Appendix, we were able to match a judge to over 95% of cases in the OGC data that we sought to match.

and 100%, however, the variation in the judge remand rate was substantial: the 25th percentile was 32% and the 75th percentile was 60%.⁴⁶³

Even as the judge-specific remand rate varies considerably, individual judges do not play an important part in explaining the variation in *district*-level remand rates. To show this, we repeat the exercise we carried out for the judge in District A. We first calculate the remand rate for each particular judge's district, and we then calculate the remand rate among all cases decided in that same district by judges *other* than the particular judge. Finally, we arrive at the particular judge's "district remand differential," as we call this measure, by subtracting the remand rate for other judges in the same district from the overall district remand rate. This measure tells us how much different the district remand rate would be if the judge in question had remanded the same fraction of cases as her district colleagues did, rather than deciding her cases as she actually did.

The following results demonstrate why we believe that individual judicial idiosyncrasy is likely to account for little of the variation in district-level remand rates. The district remand differential indicates that fewer than 2% of judges moved their district-level remand rate more than 5 percentage points in either direction: only 8 judges out of the more than 1,300 in our data include reduced their districts' remand rates by more than 5 percentage points, and only 13 judges increased their districts' remand rates by more than 5 percentage points.⁴⁶⁴ The District A judge discussed above was one of these. Only seven judges had a district remand differential in either direction of greater than 8 percentage points; three of these were in districts in which we were able to match only five cases, and the fourth accounted for all but one case we were able to match in the district in question. Ninety-five percent of all judges had a district remand

⁴⁶³ The 10th percentile was 21%, and the 90th was 71%.

⁴⁶⁴ We note that 4 of the 21 judges just described were in districts in which we were able to match judges to cases in only 5 cases.

differential of less than 2.7 percentage points in either direction, while 1,110 judges – or 81% – had a district remand differential of less than 1 percentage point in either direction.

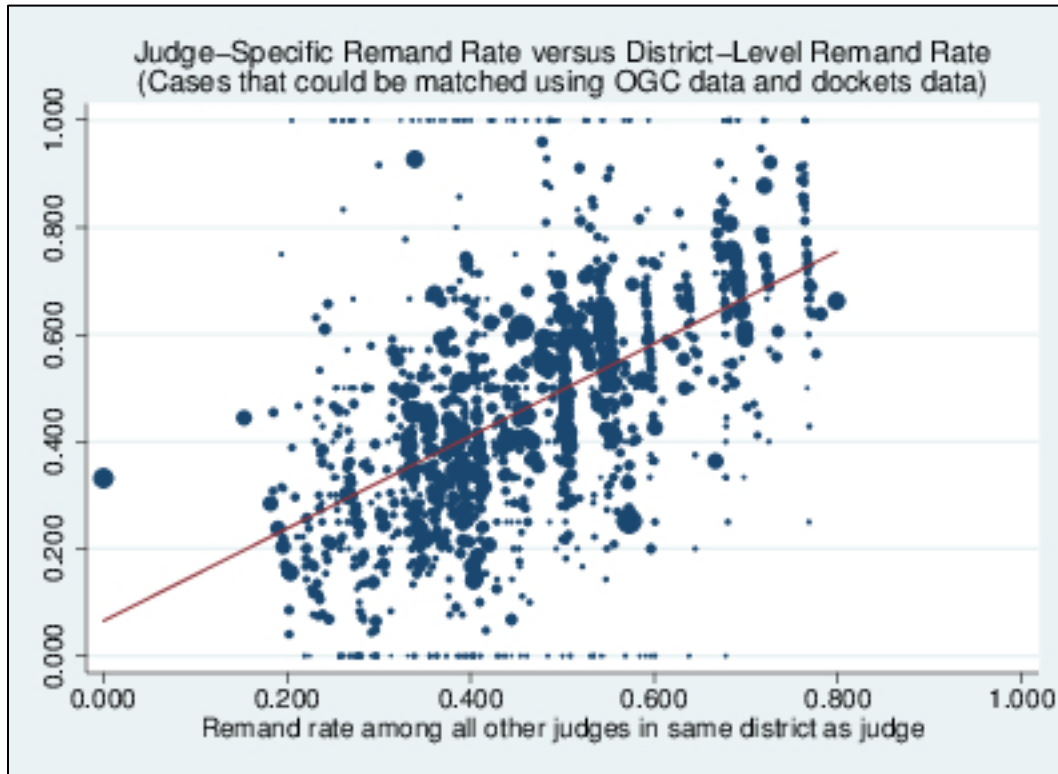
Figure 2 helps make sense of these results. It plots each judge's own remand rate on the vertical axis and the remand rate among other judges in the same district on the horizontal axis. As with the District A example above, this figure shows that judges' remand rates tend to be greater when the other judges in the same district have greater remand rates. The upwardly sloped line in the figure is the simple linear regression line.⁴⁶⁵ This line's slope of 0.863 indicates that in a district where the remand rate among a judge's peers is greater by 1 percentage point, a judge's own remand rate tends to be 0.863 percentage points greater.⁴⁶⁶ Thus when we write that judges march in lockstep rather than to their own drummer, what we mean is that judges in districts with greater remand rates tend to have similarly greater remand rates, and judges in districts with lower rates tend to have similarly low rates.⁴⁶⁷ These findings belie any suggestion that a substantial part of the variation in remand rates results from individual outlier judges.

⁴⁶⁵ In computing this line, we weighted each point by the number of cases decided by the judge in question, since judges with smaller numbers of cases contribute less information to this analysis.

⁴⁶⁶ This slope is estimated very precisely; its estimated standard error is only 0.005. Even so, visual inspection of Figure 2 makes clear that there is quite a lot of residual variation in the relationship in question. In numerical terms, the R-squared of the underlying regression estimate is 0.41, indicating that 41% of the variance in judge-specific remand rates is "explained" by variance in the remand rate for other judges in the same district. Such an R-squared value is on the high end of what one often sees in cross-sectional estimates.

⁴⁶⁷ We emphasize that we are not drawing any causal inference here. We are not, for example, claiming that judges' remand rates in a district are so highly associated because particular judges in the district influence other judges. That is a possible explanation for the high correlation between judges' own remand rates and the remand rate among all other judges in the same district. But other explanations would also be consistent with this high correlation; for example, district-level similarities in claimant characteristics or hearing office performance might explain the correlation in question.

Figure 2



3. A Circuit-Level Component to District-Level Variation

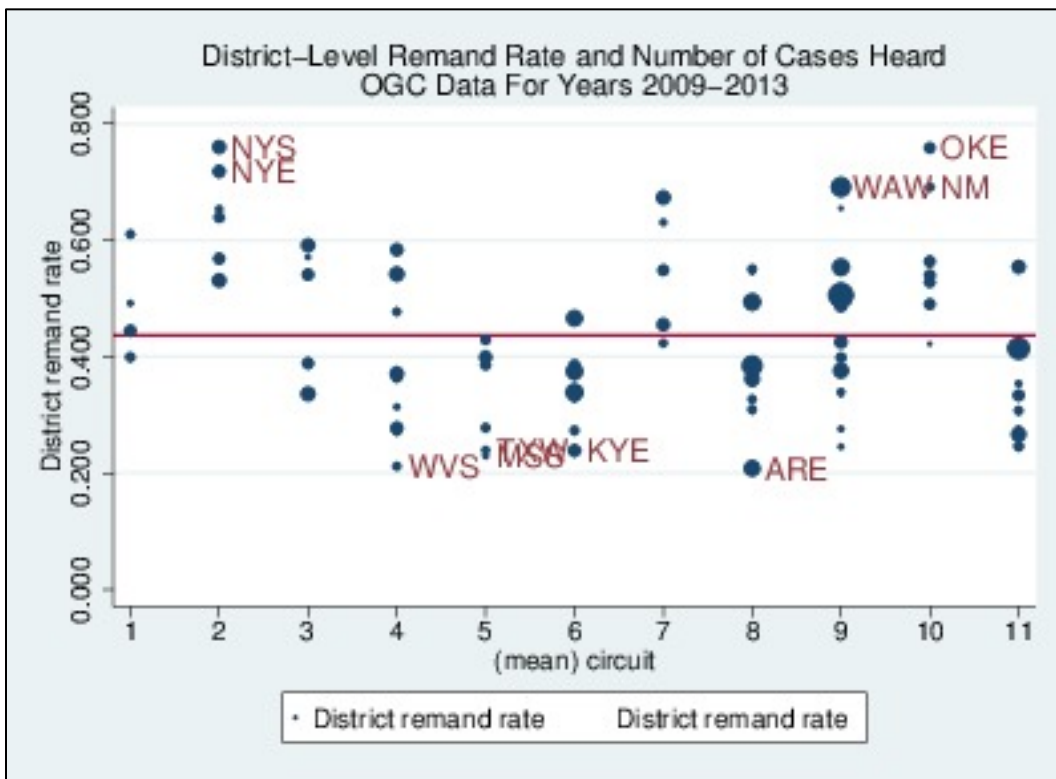
Case assignment within a district court is essentially random, so disability appeals heard by each judge within a given district should have similar characteristics. Our findings suggest that factors beyond the preferences or idiosyncrasies of individual judges determine remand rates. A district’s placement within a particular circuit is one of these. Some district-level differences result from differences in understandings of applicable law that prevail in the circuits, as well as the tone that circuits set for social security cases.⁴⁶⁸ The Seventh Circuit has earned a

⁴⁶⁸ Krent & Morris, *supra* note 1 at 396 (doubting that case law alone determines circuit-level variations, and suggesting that circuit “culture” is important as well).

reputation for its pro-claimant tilt in recent years.⁴⁶⁹ Most of the districts within it have above-average remand rates, a pattern that surely reflects the influence of Seventh Circuit decisions.

To provide a simple visual sense of circuit influence over district remand rates, Figure 3 plots the district-level remand rate on the vertical axis against the circuit to which each district belongs on the horizontal axis. Districts with more cases decided over our 2010-2013 period have circles with wider radii, indicating that they have more weight in the determination of the circuit-level remand rate. A horizontal line in the middle of the graph once again represents the district-level mean remand rate of 43.7%.

Figure 3



⁴⁶⁹ E.g., www.empirejustice.org/issue-areas/disability-benefits/litigation-legal-updates/court-decisions/seventh-circuit-decisions-too.html.

Notably, Figure 3 indicates that, for several circuits, all or nearly all districts cluster on one side of the mean. Virtually all districts in the Second, Seventh and Tenth circuits have remand rates above the mean, and most of the districts in the First Circuit also have above-mean remand rates. Nearly all the districts in the Fifth, Sixth and Eleventh Circuits have remand rates below the mean. Several circuits encompass districts with remand rates that are spread across the mean. This is so for the Third, Fourth, Eighth and Ninth Circuits. Thus, three types of circuits appear to exist: “high-remand” circuits (First, Second, Seventh and Tenth), “low-remand” circuits (Fifth, Sixth and Eleventh), and “broad-spread” circuits (Third, Fourth, Eighth and Ninth).⁴⁷⁰ The remand rates over all cases in each of these circuit groups were 36.9% for low-remand circuits, 45.0% for broad-spread circuits, and 58.2% for high-remand circuits.

An equally striking observation concerns the location of the districts with top-5 and bottom-5 remand rates. Figure 3 indicates these districts with labels to the right of the points in the figure that represent their remand rates. Four of the five top-5 districts – the Southern and Eastern Districts of New York, the District of New Mexico, and the Eastern District of Oklahoma – are located in high-remand circuits. The other one is the Western District of Washington, which is located in the Ninth Circuit. Three of the five bottom-5 districts – the Southern District of Mississippi, the Eastern District of Kentucky, and the Western District of Texas – are located in low-remand circuits. The other two – the Southern District of West Virginia and the Eastern District of Arkansas – are located in broad-spread circuits.

⁴⁷⁰ This taxonomy differs somewhat from what one might have concluded based on results Professors Krent and Morris reported. For example, they indicate that cases in the First Circuit had the lowest remand rate of any circuit. Krent & Morris, *supra* note 1 at 395. We find that the First Circuit is a high-remand circuit. We suspect this difference is the result of the fact that Krent & Morris gathered their data from Lexis searches, which, they note, might not have included all cases. *Id.* at 387. This is a commonly discussed challenge in using searches of Lexis or similar databases. See generally JOE S. CECIL ET AL., FED. JUDICIAL CTR., MOTIONS TO DISMISS FOR FAILURE TO STATE A CLAIM AFTER *IQBAL*: REPORT TO THE JUDICIAL CONFERENCE 37 n.47 (2011), at www.uscourts.gov/file/17889/download.

To measure the importance of circuit-level variation more systematically, we conducted a standard analysis of variance, focusing on circuits. This technique quantifies the share of variance in the remand rate across districts that is attributable to (i) variation across circuits and (ii) variation within circuits. The results show that cross-circuit variance accounts for 45% of the variance in district-level remand rates.⁴⁷¹ Thus, nearly half the variance in district-level remand rates is attributable to factors associated with a district's circuit. This finding comports with our sense that circuit-specific differences in the law applied to district court appeals is likely to be important. We also have no reason to doubt the conclusion that Professors Krent and Morris make in their study, that circuits can influence district court decision-making through informal mechanisms such as the tone they set.⁴⁷²

The importance of circuit influence on remand rates has implications for evaluating the significance of district-level inconsistencies and identifying who is responsible for them. To the extent that Second Circuit precedent compels a district judge in Connecticut to rule in a particular way,⁴⁷³ she cannot be blamed for an outcome that might come out differently had she been sitting in Florida and thus within the Eleventh Circuit. Courts of appeals, however, may deserve scrutiny for the degree to which they pursue (or fail to pursue) a goal of national uniformity in the elaboration of a federal legal regime. We return to this point in Part VI.

B. A Quantitative Analysis Using Multiple Regression

The results reported thus far leave a puzzle in place. Judges within a district tend to march in lockstep, and thus factors other than individual preferences explain district-level variations in the remand rate. Circuit boundaries account for a good amount of the variation, but

⁴⁷¹ The variation across circuits is statistically significant, meaning that we can reject the null hypothesis that knowing the circuit in which a district is located is irrelevant.

⁴⁷² Krent & Morris, *supra* note 1 at 397.

⁴⁷³ For an example of a district court compelled by circuit precedent to rule a certain way against its better judgment, see, e.g., *Smith v. Astrue*, Civ. No. 11-2524, 2012 WL 5269395, at *11 n.7 (D. Ariz. Oct. 24, 2012) (Campbell, J.).

a lot remains unexplained. Based on our qualitative investigation, we identified a number of possible factors. We tested these hypotheses and report our results here.

Two important caveats bear emphasis at the outset of this discussion. First, we had to make certain judgment calls in deciding which data to use for this analysis, and while we were able to use data from most district courts, we were not able to use data from all of them. The Data Appendix provides a detailed discussion of the data we use. Second, any attempt to use data on remand rates to learn about what causes one side to win or lose in litigation is fraught with problems related to litigant selection. Rational plaintiffs—or their representatives—do not embark on extended litigation journeys unless they think they are sufficiently likely to win or expect to gain something very valuable when they do win (or both). Similarly, defendants do not generally contest cases they are sure they will lose.⁴⁷⁴ Thus, the set of cases that are actually litigated is the result of deliberate choices by litigants. To the extent that claimants and/or the agency make litigation choices by reference to case characteristics or the perceived legal standards that will apply, confusing correlation with causation is hazardous.⁴⁷⁵

⁴⁷⁴ Indeed, in some cases SSA makes a request for a voluntary remand (“RVR”). The OGC data we used throughout this section do not indicate whether a remanded case involved an RVR. While the agency did provide district-level counts of RVRs by year the RVR occurred, unfortunately the OGC data do not provide federal court disposition dates. Thus the best we could do to match up district-level RVR counts with the case-level OGC data was to categorize remands for those data by the year in which the federal action was filed. When we matched the number of RVRs (based on year of disposition) to the OGC data (based on year of case filing), we found that there were a number of districts in which the RVR rate would exceed the overall remand rate, which is impossible since every RVR is a remand but not every remand is an RVR. Thus we reluctantly excluded the RVR rate from our analysis in this section. We note that the results are qualitatively similar when we do include the misaligned RVR rate described just above.

⁴⁷⁵ The classic citation on the role that selection plays in muddying the win-rate waters is George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1 (1984). For a more recent discussion tuned to the challenges of learning about changes in pleading standards from data on Rule 12(b)(6) motions, see Jonah B. Gelbach, *Locking the Doors to Discovery? Assessing the Effects of Twombly and Iqbal on Access to Discovery*, 121 YALE L. J. 2270 (2012); Jonah B. Gelbach, *The Reduced Form of Litigation Selection Models and the Plaintiff’s Win Rate* (draft on file with authors). This scholarship shows that win rate data are generally unlikely to be informative without placing some structure on the selection of cases into litigation. In the present context, there is a bit of reason for optimism, since (i) the agency generally does not consider settling cases before they are filed, and (ii) even after litigation is filed, the agency does not really “settle” in the same way that private litigants do; it either voluntarily chooses to award all benefits that would be appropriate on award or have a case remanded for further adjudication, or it seeks to have benefit denials affirmed in court. The functional impossibility of split-the-

With these caveats in mind, we now discuss the design and results of our regression analysis.

1. Hypotheses and Variables

We assessed whether district-level remand rates were associated with several types of variables:

- **District-level organization of disability appeal adjudication.** In the majority of districts, magistrate judges decide most or all social security cases. We wondered whether districts' organization and deployment of judicial resources were associated with case outcomes. If a magistrate judge decides a steady diet of social security cases to the exclusion of a more varied caseload, perhaps her decision-making tendencies more closely resemble an ALJ's. We addressed these issues by creating variables indicating whether districts assign cases mostly or exclusively to magistrates, as well as to a mix of both types of judges.
- **District-level caseload pressures.** We measured the importance of caseload pressures with two variables. The first is the share of all civil cases in the district that are disability appeals. This variable measures the relative importance of the disability part of the docket. Perhaps as the percentage of a district's civil docket devoted to social security cases grows, its remand rate falls. This would happen if judges get impatient with disability litigation and just want to clear cases off their desks.⁴⁷⁶ The second was the number of pending cases—whether civil or criminal—per congressionally approved Article III judgeship. This variable measures the overall degree of docket pressure in the district. Since the substantial evidence standard of review makes affirmances relatively easy decisions to reach, we wondered whether rising docket pressures are associated with lower remand rates.
- **Judicial skepticism of the federal government.** Professors Krent and Morris used the party affiliation of the President who nominated judges as a proxy for judges' political leanings. They determined that ideology, so measured, does not explain variation in remand outcomes.⁴⁷⁷ One district judge suggested to us, however, that judicial skepticism of the federal government, whether ideologically inflected or not, might affect decision patterns. Following her suggestion, we obtained data on the share of criminal sentences in a district in which judges sentenced criminal defendants to more lenient terms than given by federal sentencing guidelines. We included both a variable measuring a district's downward departure frequency when the U.S. Attorney sponsored the downward departure, as well as a variable measuring the frequency with which

difference deals means that what Gelbach (2012) calls “settlement selection” is not a concern here, which might make causal interpretations more plausible.

⁴⁷⁶ For empirical evidence concerning such an effect in a very different litigation context, see Eric Helland & Jonathan Klick, *The Effect of Judicial Expedience on Attorney Fees in Class Actions*, 36 J. LEG. STUD. 171 (2007).

⁴⁷⁷ Krent & Morris, *supra* note 1 at 385.

downward departures occurred in the absence of government support. We included both these variables on the theory that the government might support downward departures more often in districts where judges are more skeptical of the government as a way of channeling and thus reducing judicial rejection of the government's position.

- **District-level variables constructed in an effort to measure the performance of hearing offices within the district.** We used data from ODAR that identifies the hearing office in which newly filed court cases were initially adjudicated to create composite variables measuring the ALJ award rate from hearing offices where claims are initially adjudicated for that district, as well as the number of dispositions per day per ALJ in each hearing office.⁴⁷⁸
 - We compiled information on ALJ awards and denials from data that are publicly available on the agency's website. From this information, we estimated the hearing office-level award rate for claims filed in the hearing offices that feed into each district court.⁴⁷⁹ We speculated that higher initial award rates might be associated with lower remand rates at the district court level.
 - We also compiled information concerning the number of dispositions per day per ALJ in each hearing office.⁴⁸⁰ Because the award rate is held constant via inclusion in the regression model of the hearing office-level award rate discussed just above, dispositions per ALJ per day can be understood as a measure of labor productivity.⁴⁸¹ Including this measure allows us to determine whether more productive ALJs are associated with either higher or lower remand rates.
- **Labor market conditions in the years preceding a disability appeal.** To the extent that some workers treat disability insurance as a substitute for unemployment or a job search,

⁴⁷⁸ Ideally, we would use lagged values of our award rate and ALJ productivity variables in our analysis in order to properly match the timing of hearing office performance with subsequent district court appeals. However, data on these variables were available on the agency's website only for 2010 and later years. See www.ssa.gov/appeals/DataSets/archive/archive_data_reports.html#&ht=6. Because this means we do not have the relevant information to properly measure these variables, we used the 2010 value of each variable as a proxy for the variable's long run value in each district court.

⁴⁷⁹ See section 4(c) of the Data Appendix for the approach we took to creating judicial district-level information concerning the variables that are measured at the hearing office level.

⁴⁸⁰ Unfortunately, data we used for ALJ dispositions per day per ALJ in hearing offices, posted by the agency at https://www.ssa.gov/appeals/DataSets/archive/04_FY2010/04_September_Disposition_Per_Day_Per_ALJ_Ranking_FYTD2010.xml, do not include information for national hearing centers. Consequently, our calculation of variables determined at the hearing office level does not account for any differences across the national hearing centers and local hearing offices. However, the ODAR data we were provided indicates that national hearing centers accounted for less than 5% of cases in the filing years 2010 through 2013, so we do not believe this issue is likely to be very important.

⁴⁸¹ To be sure, this understanding may be appropriate only if ALJs have the same ratio of what are known as Type I and Type II error rates. A Type I error is a denial of a claim that in which benefits should be awarded, whereas a Type II error is an award in response to a claim that should be denied. Consider two ALJs who must adjudicate cases with the same average characteristics. If they have the same award rate and ratio of Type I and Type II errors, then on average the two ALJs are equally accurate. If one of them adjudicates more cases than the other in a typical day, then the quicker one adjudicates more cases in a given period of time, with the same average accuracy. Thus it is reasonable to describe the quicker ALJ as more productive under those conditions.

variables measuring labor market tightness might play a role in explaining initial claiming and subsequent appeal behavior. We thus included the value of the state employment to population ratio—the ratio of estimated state employment to estimated state population—in the state where each district was located. We included the employment to population ratio for the year before the year in which the case was filed in the district court, as well as for two, three, and four years before. Our thinking here was that the condition of the labor market in the several years before a disability appeal was filed in the district court should correspond roughly to its condition when claimants first decide whether to claim, then decide whether to pursue appeals within the administrative review system, and finally decide whether to appeal to federal court.

- **Salaries of lawyers.** Perhaps employment as an ALJ is particularly attractive in districts where lawyers generally earn lower salaries. If so, ALJs in low-salary areas might be particularly able lawyers. These ALJs would generate better decisions that would eventually result in lower remand rates.⁴⁸²
- **Urbanization.** Finally, we included measures of the degree of urbanization of the counties that make up each district. One interview subject suggested that federal judges in more rural districts probably remand fewer cases.⁴⁸³

The data we used for the regression analysis below ultimately included information on eighty-three districts. Collecting data on all the variables described above was not always simple, and in some instances we had to make do with rough proxies. The Data Appendix includes more details.⁴⁸⁴

2. Statistical Model and Results

To assess the importance of the variables described above, we used the OGC data to calculate annual district-level remand rates for the years 2010, 2011, 2012, and 2013. We then employed a two-step statistical approach to examine the factors associated with the annual district-level remand rate.⁴⁸⁵ In Step 1, we used ordinary least squares (“OLS”) to estimate a linear regression relating the annual district-level remand rate to district-level values of (i) a

⁴⁸² The only nationwide data on lawyers’ salaries we could find were for 2014. See Data Appendix.

⁴⁸³ Federal Judge 24 at 8.

⁴⁸⁴ We note that our results were not qualitatively different when we included our best effort at measuring RVR rates (see note 474, *supra*); nor were they different when we included a variable measuring the share of disability appeal cases in which the docket indicated *pro se* litigation (we used our court dockets data to construct this variable).

⁴⁸⁵ We used the Stata module “fese” to conduct this analysis. See Austin Nichols, *fese: Stata module calculating standard errors for fixed effects* (2008), at <http://ideas.repec.org/c/boc/bocode/s456914.html>.

number of time-varying variables and (ii) dummy variables indicating the district court with which each observation was associated.⁴⁸⁶ The time-varying variables are:

- the share of decisions that awarded benefits in hearing offices associated with a district in the year in question;
- the number of overall cases—civil or criminal—pending per judgeship in the district in the year in question;
- the ratio of disability appeals filed to all civil cases filed in the district in the year in question;
- for each district and year, the share of criminal sentences in which a downward departure from sentencing guidelines occurred where (i) the government supported the departure, as well as (ii) the share of sentences in which the judge made a downward departure or other below-range sentence without government support;
- three dummy variables indicating whether the case was filed in the district court in 2011, 2012, or 2013 (2010 is the reference year).

Step 2 of our statistical procedure is built around the estimated coefficients on the district-level dummies. These coefficients may be thought of as the district-level remand rate that would have been expected in 2010 if, within each district, the time-varying variables assumed their district-level average values for the period.^{487,488} We refer to each of these values as the “adjusted district-level remand rate” for the period 2010-2013. We then estimated a linear regression relating the adjusted district-level remand rate to those of our explanatory variables that vary across district (or state) but do not vary over time. These are variables whose values are (i) different in, say, the Eastern District of New York and the District of Connecticut, but

⁴⁸⁶ An estimator that includes such dummy variables is sometimes referred to as the “fixed effects” estimator, in this instance allowing for district-level fixed effects.

⁴⁸⁷ This is a well-known fact about fixed effects estimates, i.e., the estimated coefficients on dummies indicating category membership.

⁴⁸⁸ To get each district’s corresponding value for the years 2011, 2012, or 2013, one would just add the estimated coefficient on the corresponding year’s dummy variable to each estimated district coefficient. Adding the same number to each of these district coefficients would not change the coefficients estimated in our second step (except for the reference category, which does not play a substantive role in our analysis).

which are (ii) the same every year within the Eastern District of New York and within the District of Connecticut.⁴⁸⁹ These time-constant variables are:

- the average number of daily dispositions per ALJ in hearing offices associated with a district in 2010;
- the share of decisions that awarded benefits in hearing offices associated with a district in 2010;
- dummy variables indicating whether a district's disability appeals are primarily handled by magistrates, by Article III judges, or by a mix;
- state-level median lawyer salaries for 2014; and
- variables measuring the degree of urbanization of the counties that make up each district.

⁴⁸⁹ For example, one such variable is the degree of urbanization of counties within the district. This variable obviously varies across districts. But since it is measured using data from a single decennial census, it does not vary over time within a given district.

a. Regression results for the first step

The following table presents results from the first step, for the time-varying variables:

Table 2: First-Step Remand Regression Results

Dependent variable: Share of Cases Remanded	Coefficient	Standard Error	<i>t</i> -ratio
<i>Employment-to-population ratio:</i>			
One year before year case filed	21.6	14.8	1.46
Two years before year case filed	-12.0	12.2	-0.98
Three years before year case filed	22.0	15.2	1.44
Four years before year case filed	-24.5	10.9	-2.25
<i>District caseload variables</i>			
Number of cases pending per judgeship	0.5	0.2	3.03
Disability appeals' share of civil cases	-0.9	1.2	-0.76
<i>Downward departures</i>			
Government sponsored	-3.6	2.2	-1.65
Other	-2.2	3.1	-0.72
<i>Year effects and constant</i>			
2011 difference, relative to 2010	-0.3	4.2	-0.07
2012 difference, relative to 2011	2.7	6.1	0.44
2013 difference, relative to 2012	-5.7	5.1	-1.13

Notes: All non-dummy variable explanatory variables were scaled so that a one unit change is expressed in interquartile range units, i.e., corresponds to a move from the 25th to the 75th percentile of the variable in question. Standard errors were computed using cluster-robust variance estimation, clustering on state. The intercept, which corresponds to the average value of the estimated district-level dummy coefficients with all other variables held at their means, was 44.8.

Only two of the first-step explanatory variables were statistically significantly different from zero (significant differences are denoted with bold-font display for the rows corresponding to each variable). The employment-to-population ratio in the year four years before a claimant files the disability appeal in federal court is negative and statistically significant. One possible explanation for this result begins with the observation that a high employment-to-population ratio generally indicates a strong labor market, one that presents workers with good job options.

With a stronger labor market, people on the margin between work and claiming disability benefits will keep working. We can imagine implications for the district's remand rate. If such people would be less likely to qualify for benefits, then their self-selection out of the pool of claimants should improve the general strength of claims that do get filed. Assuming that strong claims are easier for the agency to decide correctly, the four layers of agency review would get the results right more often and leave fewer weak agency decisions for federal courts to vacate.

The magnitude of the coefficient on the four-year lag of the employment-to-population ratio is substantial. The explanatory variable itself is scaled so that a one-unit change corresponds to a move from the 25th percentile of this variable to the 75th percentile—here, a move from a four-year lag of the employment-to-population ratio of 42.5% to one of 46.7%. The coefficient of -24.5 on the four-year lag of the employment to population ratio means that an increase of 4 percentage points in the employment-to-population ratio is associated with a decline in the remand rate of about 25 percentage points – obviously a very large effect. We discuss the significance of this finding in Section C.

We note that the third lag of the employment-to-population ratio, while statistically insignificant, is nearly as large in magnitude as the coefficient we have just discussed, but with the opposite sign. The same is true of the first lag. Ignoring the statistical insignificance of these estimates for the sake of argument, we could imagine explanations for these effects.⁴⁹⁰ However, given their statistical insignificance, we will not dwell on them.

⁴⁹⁰ For instance, perhaps ALJs and those conducting appellate review within the agency are more likely to believe that a claimant could find appropriate work when the economy is strong. If so, then claims that should be allowed would be more likely to be denied at times when the economy is strong. Subsequently elevated rates of remand would then occur at the district court level. The agency has told us that its policy is not to take into account the strength of the economy in adjudicating appeals and that adjudicators do not in fact do so. Of course, a policy's presence does not guarantee blanket compliance with that policy (as illustrated, for example, by the agency's own laudable efforts to improve the policy compliance of its internal decision making).

The other time-varying factor that is statistically significantly associated with the remand rate is the district's overall caseload, as measured by the ratio of pending cases, whether civil or criminal, to the number of judgeships approved for the district. However, the estimated coefficient for this variable is quite small. A move from 350 cases pending per judgeship to roughly 550 –corresponding to a move from roughly the 25th to the 75th percentile – is associated with an increase in the remand rate of only 0.5 percentage points. Even assuming that higher caseloads in fact push remands up, it would take a change in the district's caseload of roughly a thousand cases to move the remand rate by just 2 percentage points. Thus, we conclude that caseloads are a statistically significant correlate of remand rates, but not one that seems particularly important in substantive terms.

b. Regression results for the second step

Our second-step estimates capture the association between the time-constant variables, on the one hand, and the district-level fixed effects from the first step estimation, on the other. Recall that a district's estimated fixed effect may be interpreted as the remand rate across that which would have been expected to prevail in 2010 had all time-varying factors been held at their district-level averages over this period. Results from the second step appear in Table 3.

Table 3: Second-Step Estimates — Association Between First-Step Estimated District Fixed Effects and Time-Constant Variables

Dependent variable: district-level estimated fixed effect from first step	Coefficient	Standard Error	<i>t</i> -ratio
<i>Hearing Office variables:</i>			
Number of dispositions per ALJ per day	0.4	2.5	0.17
Award rate	-3.8	2.5	-1.53
<i>Assignment of cases within district (reference category is “Mostly to Article III judges”):</i>			
Mostly to magistrates	2.0	4.2	0.48
To a mix	-4.4	6.2	-0.72
Information not available to us	-13.1	7.8	-1.68
<i>Urbanicity of counties in district:</i>			
Fraction of residents living in counties with 1 million or more residents	-6.2	4.7	-1.33
Fraction of residents living in counties with 20,000 or fewer residents	-6.7	2.6	-2.61
<i>Lawyers’ outside options:</i>			
Median salary of lawyers in state, 2014	-3.4	2.6	-1.29
<i>Circuit effects (reference category is Eleventh Circuit):</i>			
First	13.8	7.9	1.74
Second	27.3	7.4	3.7
Third	24.5	12.1	2.02
Fourth	8.2	6.2	1.32
Fifth	-8.9	6.1	-1.46
Sixth	0.4	6.1	0.06
Seventh	12.9	5.1	2.54
Eighth	-6.3	6.6	-0.95
Ninth	18.6	6.9	2.7
Tenth	16.8	7.3	2.3
Reference category	37.6	6.8	N/A

Notes: All non-dummy variable explanatory variables were scaled so that a one unit change is expressed in interquartile range units, i.e., corresponds to a move from the 25th to the 75th percentile of the variable in question. Standard errors were computed using cluster-robust variance estimation, clustering on state.

Reference category coefficient should be interpreted as the 2010 remand rate expected in a district in the Eleventh Circuit in which most cases are assigned to Article III judges, with all non-dummy time-varying and time-constant variables taking on mean values.

Our findings from this second step are easy to summarize:

- Only one variable other than the circuit dummies had a coefficient that was statistically significantly different from zero. This variable, the fraction of residents in counties encompassed by the district who live in counties with fewer than 20,000 residents, is associated with a relatively substantially lower remand rate.⁴⁹¹
- There is no statistically significant relationship between district remand rates and our measures of ALJ productivity for 2010. At least as can be measured with these variables, then, the strength of federal court appeals is not meaningfully associated with the number of cases ALJs decide.
- Neither the assignment of cases across magistrates/Article III judges nor outside options for lawyers (as best as we can measure them using state-level salary information for 2014) are associated with district-level remand rates.

To understand the results for the circuit dummies, it is important to understand that the reference category in Table 3 is the Eleventh Circuit.⁴⁹² The coefficients on the First through Tenth Circuit dummies are expressed in terms of deviations from the Eleventh Circuit. Other things equal, then, we would expect a remand rate 24.5 percentage points greater in the Third Circuit than in the Eleventh. For the Second, Third, Seventh, Ninth and Tenth Circuits, we would expect remand rates between 13.2 and 27.3 points above the Eleventh Circuit's. For the

⁴⁹¹ Moving from a value of this variable of 3% to 24%, which corresponds to a move from the 25th percentile to the 75th percentile, is associated with a drop of about 7 percentage points in the remand rate.

⁴⁹² The reference category coefficient of 37.6 tells us that, if all non-circuit dummy variables were held at their means, then the average remand rate for districts in the Eleventh Circuit in 2010 would be expected to be 37.6% (assuming that cases were handled by mostly Article III judges, which is the reference category for the "Assignment of cases within district" variables in Table 3).

ODAR has pointed out to us that under SSA's Acquiescence Ruling 99-4(11), Appeals Council dismissals of requests for review from ALJ decisions are handled differently when the claimant resides in Alabama, Florida or Georgia, than when the claimant resides outside the Eleventh Circuit. Within the Eleventh Circuit states, such dismissals are reviewable by district courts, pursuant to AR 99-4(11) following *Bloodsworth v. Heckler*, 703 F.2d 1233 (11th Cir. 1983); outside these states, such dismissals are non-reviewable, pursuant to 20 CFR §§ 404.972, 416.1472. It is of course possible that the reviewability of dismissals in the Eleventh Circuit, and no other, creates differences in the case mix between the Eleventh Circuit and other circuits. However, we are unaware of any empirical evidence that such differences exist. Nothing of substance would change if we used a different circuit as the reference category, in any case, since doing so would just move all the circuit dummy coefficients up or down by the amount of the coefficient on the new reference circuit.

other circuits, the estimated circuit coefficients are statistically insignificant and also smaller in magnitude.

We note that the circuit results correspond quite well to the classification of circuits we offered above. The Eleventh Circuit is among the low-remand circuits according to this classification, while the high-remand circuits were the First, Second, Seventh, and Tenth Circuits. Table 3 displays a significant circuit coefficient for all of these but the First Circuit. The Third and Ninth Circuits, which have significantly positive circuit coefficients in Table 3 were among our broad-spread circuits according to the classification above.

Taken together, the results from our regression analysis above suggest several basic conclusions. First, with the exception of a single variable related to rural population share, the factors we were able to quantify concerning conditions in the district courts themselves do not appear to explain the variation in district-level remand rates. Second, the same is true for quantifiable factors related to ALJ productivity and ALJ allowance rates, as best as we can measure them. Third, labor market conditions that can be expected to affect the quality of initial claims are strongly associated with remand rates in the district courts several years down the road. Fourth, there is inarguable evidence of substantial circuit-level differences in remand rates.

C. Using ODAR Regions to Capture Intra-Agency Factors

We can confidently conclude that circuit boundaries are a significant factor in producing differences among district remand rates. But none of our variables, save one, did much to explain the residual variation that circuit influence does not produce. The one variable that does matter, the strength of the state labor market several years before appeals are filed in the district court, points to a potential determinant more interesting, but also harder to detect, than circuit boundaries. In theory, regional differences in labor market conditions should not alter the makeup of district courts' social security dockets. Claims pass through four layers of review

within the agency. At each step, the agency's policy is to measure claims against a single national standard, with some variation here and there prompted by acquiescence rulings. If this review were sorting among claimants with equal success nationwide, then regional differences in claim quality generated by labor market variation should be eliminated by the time claims make their way to the end of the agency's review process. In theory, then, labor market variation should have no impact on the quality composition of cases left for claimants to file with the federal courts at the end of the administrative review process.

Our intuition, however, is that internal agency review of claims is unlikely to be uniform across the country. Some hearing offices likely generate better decisions than others. Analysts in one ODAR division might review appeals slightly differently than others. If we are right, there could still be regional differences in the quality of inputs for district court decision-making, differences that might color judicial impressions of ALJ decisions district-wide. For example, Division II of the Appeals Council reviews appeals taken from the Pasadena and San Jose Hearing Offices. If Pasadena ALJs happen to generate weaker decisions on average than their colleagues in San Jose, but if the Appeals Council remands decisions from the two hearing offices at about the same rate, then a potential appeal to the Central District of California from a Pasadena ALJ will be stronger on average than an appeal to the Northern District of California from a San Jose ALJ.

If judges see enough such differences, they might develop expectations related to the quality of agency review that vary across hearing offices. The same process could generate variation across judicial districts in judges' attitudes toward agency review in general. This story of agency-judiciary interaction thus would be able to explain why variation in remand rates does not seem to be much driven by individual judges, but does vary considerably at the district level.

To test such hypotheses systematically would require substantial quantitative data on the agency's internal review process, and we do not have such data. However, we did identify an approach that allowed us to scratch the surface of this issue. Appeals from denials of claims by state DDS offices are channeled into ten ODAR regions.⁴⁹³ ODAR's website provides a list of the hearing offices in each region.⁴⁹⁴ These regions mostly overlap the circuit boundaries, but with some differences. ODAR Region 8, for example, includes Utah (a Tenth Circuit state), Montana (a Ninth Circuit state), and North Dakota (an Eighth Circuit state). We are thus able to augment the second-step of our estimation method above by adding dummies that indicate which ODAR region a case would have been handled by. This approach allows us to investigate whether district-level remand rates vary systematically across ODAR regions, even while controlling for differences in the circuits via the continued inclusion of circuit dummies.⁴⁹⁵

To understand the logic of this approach, consider cases filed in North Carolina and South Carolina. These states are both assigned to the Court of Appeals for the Fourth Circuit, and they are also both assigned to ODAR's Region 4. Other states in ODAR Region 4 are assigned to other circuit courts, so comparing remand rates for the Carolinas to those for the

⁴⁹³ There are also five national hearing centers that hear cases from hearing offices around the country. However, these account for relatively few cases. For example, data posted on the agency's website show that between September 26, 2015, and March 26, 2016, there were 311,785 dispositions across all hearing offices. Of these, the National Hearing Centers in Albuquerque, Baltimore, Chicago, Falls Church, and St. Louis accounted for 11,812—or less than 4 percent. Appeals can also be from partially favorable decisions, and they can also be from favorable decisions if the claimant contests the onset date.

⁴⁹⁴ See, e.g., https://ssa.gov/appeals/odar_ho_sites.html (listing hearing offices by region); see also https://www.ssa.gov/appeals/ODAR_Hearing_Sites_Map.html#&sb=0 (providing a map of hearing offices and regions). We note that the agency has stressed to us that not every appeal is necessarily handled by personnel in the ODAR region corresponding to the state where the underlying claim originated. For example, a small percentage of cases are heard in national hearing centers. See, e.g., *supra* note 493. And the agency has told us that “some hearing offices have servicing areas that overlap into different states.” It provided the example of the Cincinnati hearing office, which “handled cases in Ohio, Indiana, and Kentucky – three different states, two different federal circuits, and two different regional offices.” We do not, however, understand that there are many cases affected by such deviations from the otherwise straightforward correspondence between hearing offices, states, and ODAR regions provided in the ODAR webpages cited above.

⁴⁹⁵ The agency stressed to us that cases get moved around within ODAR and that ODAR has national hearing centers outside the regional structure. Still, we believe the approach described here may capture systematic differences within the regional structure.

other ODAR Region 4 states helps identify the remand rate component associated with assignment to the Fourth Circuit.⁴⁹⁶ Other states in the Fourth Circuit are assigned to different ODAR regions, so comparing remand rates for the Carolinas to this second set of other states helps identify the remand rate component associated with assignment to ODAR Region 4.⁴⁹⁷ By including both circuit and ODAR region dummies in our second step, we are able to separate out these two components of variation in district-level remand rates.

Table 4 reports the coefficients on the circuit dummies and ODAR region dummies from a revised second-step estimation that added the ODAR region dummies to the set of explanatory variables.⁴⁹⁸ As before, the Eleventh Circuit is the reference category. The circuit coefficients are again positive and significant for the Second, Third, Seventh, and Tenth Circuits. For the Seventh and Tenth Circuits the coefficients' magnitudes are roughly equivalent to their corresponding values from the original results displayed in Table 3. However, the magnitudes for the Second and Third Circuit coefficients are much greater – 49.6 percentage points and 37.9 percentage points, respectively – with the ODAR region dummies included. Further, circuit coefficients are positive, statistically significant, and substantial in magnitude – 40.1 percentage points and 15.0 percentage points, respectively – for the First and Fourth Circuits. Taken together, these modified circuit results are notable. They show that when the ODAR region is accounted for, district-level remand rates have large circuit-specific components not only in all circuits we previously classified as high-remand on the basis of raw district-level remand rates (First, Second, Seventh, and Tenth) but also in two of the broad-spread circuits (the Third and Fourth Circuits).

⁴⁹⁶ The other Region 4 states are Alabama, Florida, and Georgia (all in the Eleventh Circuit); Kentucky and Tennessee (both in the Sixth Circuit); and Mississippi (in the Fifth Circuit).

⁴⁹⁷ The other Fourth Circuit states—Maryland, Virginia and West Virginia—are all in ODAR Region 3.

⁴⁹⁸ Because we generally found qualitatively similar results for the coefficients on the other variables included in the second-step when we included the ODAR region dummies, we do not display these coefficients in Table 4.

Table 4: Second-Step Estimates — Association Between First-Step Estimated District Fixed Effects and Time-Constant Variables

Dependent variable: district-level estimated fixed effect from first step	Coefficient	Standard Error	<i>t</i> -ratio
<i>Circuit effects (reference category is Eleventh Circuit):</i>			
First	40.1	13.5	2.96
Second	49.6	14.8	3.36
Third	37.9	10.7	3.54
Fourth	15.0	6.6	2.27
Fifth	-9.5	6.0	-1.58
Sixth	4.7	8.2	0.58
Seventh	17.7	8.0	2.22
Eighth	-1.3	8.1	-0.16
Ninth	11.8	16.1	0.73
Tenth	22.2	8.9	2.49
<i>ODAR region effects (reference category is region 4):</i>			
1	-27.3	12.1	-2.26
2	-16.4	11.6	-1.41
3	-25.4	5.5	-4.6
5	-5.3	6.0	-0.88
6	-0.3	4.5	-0.07
7	-6.1	8.6	-0.71
8	-4.5	10.2	-0.44
9	1.8	16.9	0.11
10	24.6	16.3	1.51
Reference category value	36.7	7.5	4.92

Notes: All non-dummy variable explanatory variables were scaled so that a one unit change is expressed in interquartile range units, i.e., corresponds to a move from the 25th to the 75th percentile of the variable in question. Standard errors computed using cluster-robust variance estimation, clustering on state, to account for unmodeled spatial correlation within states.

The estimated coefficients on the ODAR region dummies themselves are equally intriguing. The ODAR region reference category is Region 4, so we have estimated coefficients

for regions 1-3 and 5-10.⁴⁹⁹ Districts in Regions 5-9 have coefficient estimates that are both small in magnitude and statistically insignificant; Region 10 has a healthy coefficient of 24.6 percentage points, but this estimate is very imprecise and is statistically insignificant. The main story in Table 4 is the set of coefficient estimates for Regions 1 and 3.⁵⁰⁰ These estimates tell us that for districts in Regions 1 and 3, remand rates are statistically significantly *below* remand rates in Region 4. Moreover, the differences are sizable—roughly 25 percentage points.

To understand these results, consider a comparison of the District of Connecticut with reference-category districts, those in Region 4 and the Eleventh Circuit. As the last row in Table 4 shows, our augmented second-step model has a reference category remand rate of 36.7%. Because Connecticut is in the Second Circuit, we would expect its remand rate to be 49.6 percentage points greater than the remand rate in districts in the Eleventh Circuit. Since Connecticut is also in Region 1, whose second-step regression coefficient is -27.3, Connecticut's district-level remand rate is also expected to be 27.3 percentage points below the level for districts in Region 4. The overall difference between the District of Connecticut and reference-category districts, then, is found by combining the Second Circuit coefficient, which adds 49.6 percentage points, with the Region 1 coefficient, which subtracts 27.3 percentage points. The net difference is thus 22.3 ($= 49.6 - 27.3$) percentage points.

In other words, this model's results tell us we should expect Connecticut to have a remand rate that is roughly 22.3 percentage points greater than reference category districts (all else equal). This net impact is almost identical to the result we would get from comparing Second Circuit districts—including the District of Connecticut—to districts in the Eleventh Circuit using our earlier results in Table 3 (the Second Circuit's coefficient there is 27.3

⁴⁹⁹ Thus the reference category consists of districts that are in ODAR Region 4 and in the Eleventh Circuit. These are districts in Alabama, Florida, and Georgia.

⁵⁰⁰ Region 2 also has a sizable negative coefficient, -16.4 percentage points, but it is statistically insignificant.

percentage points, just a few points above the 22.3-point figure for Table 4). But our Table 4 results show that the path to this conclusion matters a lot. When we allow remand rates to differ across ODAR regions within circuits, we find that Connecticut's membership in the Second Circuit is associated with an enormous increase in the expected remand rate, while its membership in Region 1 is associated with a smaller but still very sizable reduction in the expected remand rate.⁵⁰¹

This constellation of results suggests that something important is happening not just within districts in the same circuit, but also within districts for which appeals are handled in Regions 1 and 3.⁵⁰² The statistical approach we have taken here cannot reveal what those important things are. They could include systematic differences at the state level, including, for example, variation in the performance of state DDS determinations that are not eliminated by the agency's own processes. Whatever the reasons, it seems very possible to us that features of the agency's review processes, and possibly also its initial claims determination processes, play a real role in causing variation in district-level remand rates.

If the quality of inputs for district court decision-making were identical across the country, and if federal judicial attitudes toward social security cases were also uniform, then circuit boundaries should generate most, if not all, of the differences in district remand rates. But we have already seen that this is not so, and Table 4 helps us generate some thoughts about why.

One explanation regards inputs—case quality and decisional quality—as uniform and attributes residual variations to differences in federal judicial attitudes motivated by factors unrelated to anything the agency is doing. Neither we nor Professors Krent and Morris were able

⁵⁰¹ The agency has noted to us that ODAR Region 1 follows different regulations from those other regions follow. Just as differences in circuit law provide a possible explanation for cross-circuit variation in remand rates, such differences in regulations provide a possible explanation for differences across ODAR regions.

⁵⁰² We note that the coefficients for Regions 2 and 10 are large in magnitude, though neither was statistically significant.

to find any such factors. We also find it unlikely that judicial attitudes would coincidentally correspond to ODAR region boundaries.

Another explanation chalks up some of the residual variation to geographic variations in case quality. Hypothetically, cases appealed to the District of Colorado might be stronger on average than those appealed to the Western District of Missouri even if decision-making within the agency is not uneven and even if the pools of potential appeals in the two districts are equally strong. Claimant representatives in Denver might simply be better at picking cases than their colleagues in Kansas City. We lack data to test this possibility, but we doubt that it is right. As before, the correspondence of lawyer quality to ODAR region boundaries that would be required for such an explanation strikes us as implausibly coincidental. We continue to believe that decision-making within the agency is uneven in places. Our in-depth qualitative investigation into three districts and the hearing offices within them support this intuition.

D. Qualitative Evidence of Varying Input Quality

Before selecting variables for quantitative testing, we examined three districts closely in order to determine what factors might produce their district-level disparities. In one district the agency prevails in most cases (the “low remand district”), and in another claimants prevail in most cases (the “high remand district”). The third district had an average remand rate (the “average remand district”). The three districts contrasted in other interesting ways.⁵⁰³ We interviewed district and magistrate judges in these districts, spoke with their law clerks, and talked to government and private lawyers who litigate in them. In addition, we interviewed ALJs

⁵⁰³ For example, one district was almost entirely urban and densely populated. Another was mostly rural. The third included both rural and metropolitan areas.

and decision writers who work in some of the hearing offices that generate decisions that, if appealed, would likely go to these districts.⁵⁰⁴

Interviews revealed striking, obvious differences among the districts. The average remand district came off as just that – average in almost every respect. For this reason, we do not describe in any detail the results of our interviews of personnel who work within it. In contrast, a deep strain of discontent was palpable in our interviews with high remand district personnel. The agency personnel and federal judges in the low remand district, in contrast, expressed a much more positive view of disability claims adjudication and litigation. We summarize these findings in this section.

We do not identify the exemplar districts by name, as we appreciate that the agency might have legitimate concerns with the use of any of them as a model of institutional health or dysfunction.⁵⁰⁵ We expect that each federal district and the agency offices within them differ in particular ways, and our interview data are not robust enough to offer more than an impressionistic sense of what goes on in hearing offices and chambers. These differences, though, are the point. We describe some of the ways in which the high and low remand districts compare and contrast to explain the basis for our hypothesis that varying input quality explains some of the inconsistencies in district court remand rates. Our claim is a simple one: within an agency administering a national program, different offices contrast in ways that might have implications for judicial confidence in agency decisions.

1. Contrasts in the Hearing Offices

⁵⁰⁴ A claimant files for review in the district in which she lives or has her principal place of business. *See* www.socialsecurity.gov/appeals/court_process.html. “To the extent possible,” hearings are conducted within seventy-five miles of a claimant’s home. HALLEX, *supra* note 38 § I-2-3-10. Sometimes claimants appeal from decisions rendered in National Hearing Centers. *See infra* note 110 & accompanying text. Although these centers do not take cases from different hearing offices at equal rates, the overall percentage is small enough such that we are comfortable with our description.

⁵⁰⁵ The agency is aware of these districts’ identities.

The hearing office that we studied in the low remand district came off as a model of institutional health. “I can’t begin to think of a better place to work,” one decision writer told us.⁵⁰⁶ The office has stable management, with a long-serving Hearing Office Chief Administrative Law Judge (“HOCALJ”),⁵⁰⁷ and ALJs tend to stay once they are assigned there.⁵⁰⁸ ALJs consistently expressed their high regard for the quality of in-house decision writers. “We have fabulous writers here,” one ALJ told us, expressing a common sentiment.⁵⁰⁹ Each ALJ we interviewed preferred keeping decision-writing local, as opposed to sending decisions to the national case assistance centers for writing.⁵¹⁰ Familiarity with the region, the circuit, and individual ALJs and their idiosyncrasies made local writing “so much better,” one ALJ told us.⁵¹¹ Personnel also described collegial, easy communication between ALJs and decision writers.⁵¹²

Another striking theme from our interviews with hearing office personnel in the low remand district was their regard for district court feedback. Several ALJs and decision writers complained of nit-picky remands⁵¹³ and of judicial ignorance of the agency decision-making process.⁵¹⁴ A number conceded, however, that district court remands are justified and provide useful instruction.⁵¹⁵ The hearing office has formalized this instruction in semi-annual

⁵⁰⁶ Decision Writer 3 at 1.

⁵⁰⁷ ALJ 10 at 1.

⁵⁰⁸ ALJ 9 at 3; ALJ 10 at 4.

⁵⁰⁹ ALJ 18 at 4; *see also* ALJ 23 at 2 (insisting that seven or eight of the decision writers are as good as any in the country); ALJ 9 at 2-3 (“Writing in this office is exceptional . . . I love the writers here – they are on a whole other level.”); ALJ 15 at 3 (insisting that writers are “really, really good”); ALJ 10 at 2 (“We’ve got some good writers.”); ALJ 17 at 3 (“We’ve got some of the best decision writers in the country.”); ALJ 19 at 2 (“They are great.”).

⁵¹⁰ *E.g.*, ALJ 23 at 2; ALJ 17 at 3; ALJ 19 at 2.

⁵¹¹ ALJ 18 at 3.

⁵¹² ALJ 23 at 2; ALJ 9 at 3; ALJ 15 at 3; ALJ 19 at 2; Decision Writer 6 at 2; Decision Writer 3 at 2; Decision Writer 9 at 2.

⁵¹³ ALJ 15 at 3; ALJ 18 at 4; ALJ 23 at 2; Decision Writer 9 at 4; Decision Writer 3 at 3.

⁵¹⁴ ALJ 18 at 4.

⁵¹⁵ ALJ 9 at 3; ALJ 15 at 3; ALJ 17 at 2 (some remands are “absurd” but “most” are “really good”); ALJ 23 at 2 (most remands are well-reasoned); *see also* Decision Writer 9 at 4; Decision Writer 6 at 3 (explaining that decision writers and ALJs “care about [district court] holdings”).

memoranda, circulated to all ALJs and decision writers, that summarize social security decisions the district court issues.⁵¹⁶ This practice seems popular, at least with the ALJs we interviewed.⁵¹⁷ Several reported that it has improved their decision-making.⁵¹⁸

Most of the hearing office personnel from the high remand district described a very different and more problematic work environment.⁵¹⁹ Several ALJs complained of poor quality decision writing,⁵²⁰ and several expressed a preference for decisions written off-site in national case assistance centers.⁵²¹ An ALJ described unstable, volatile management at a hearing office for much of the past decade,⁵²² and ALJs and a claimant representative complained of the office's capacity to perform basic administrative tasks.⁵²³ Some personnel described communication difficulties between ALJs and decision writers.⁵²⁴

Attitudes regarding judicial review in the high remand district were mostly negative. Like their low-remand counterparts, ALJs here described remands as “nit-picky”⁵²⁵ and complained that district judges have little understanding of or regard for agency processes.⁵²⁶

⁵¹⁶ ALJ 10 at 3; Decision Writer 9 at 4.

⁵¹⁷ ALJ 9 at 3; *see also* ALJ 15 at 3 (responding “very much so” when asked if memoranda on case law are helpful).

⁵¹⁸ ALJ 10 at 3; ALJ 9 at 4; ALJ 23 at 3.

⁵¹⁹ ALJ 20 at 6; ALJ 3 at 2; ALJ 4 at 2, 4; ALJ 5 at 2.

⁵²⁰ ALJ 20 at 2 (suggesting that decision writers in other hearing offices are “miles ahead” of what some decision writers in this ALJ’s hearing office produce); ALJ 16 at 1 (describing some decision writers as “greatly lacking”); ALJ 16 at 3 (estimating that “30-40%” of decision writers have “performance issues”); *id.* at 2 (expressing concerns over some decision writers’ productivity); ALJ 5 at 4 (ALJ describing decision writing as “hit or miss”); *id.* at 4 (commenting that the hearing office has “no shortage” of writers who are “less than strong”); ALJ 20 at 2 (ALJ describing the “best” writing as “mediocre”).

⁵²¹ ALJ 20 at 2; ALJ 16 at 2; ALJ 3 at 2. *But cf.* ALJ 4 at 4 (complaining of communication difficulties with off-site writers); ALJ 6 at 3 (expressing indifference).

⁵²² ALJ 3 at 2; ALJ 4 at 2.

⁵²³ One ALJ expressed “absolutely no faith in any of the management initiatives” being tried in the hearing office. ALJ 4 at 3. Another ALJ complained that employees routinely fail to show up for work. ALJ 16 at 2. A third ALJ suggested that support staff were overmatched by the work they were supposed to perform. ALJ 3 at 2. A claimant representative who practices in the hearing office described difficulties sending in faxes or getting phone calls answered at what this lawyer labeled a “disorganized” office. Claimant Representative 4 at 2.

⁵²⁴ ALJ 3 at 2; ALJ 4 at 4; ALJ 20 at 5; ALJ 3 at 4; Decision Writer 11 at 2. *But see* Decision Writer 4 at 2 (claim by a decision writer that he can discuss issues with ALJs directly).

⁵²⁵ ALJ 20 at 3; ALJ 4 at 5.

⁵²⁶ An ALJ complained that federal judges have unreasonable record development expectations. ALJ 4 at 1, 5. Another ALJ complained that district judges don’t understand that she lacks time to put citations in her instructions and to edit decision drafts. ALJ 20 at 3. One ALJ told us that “sometimes” a district court catches her mistake, but

But none expressed countervailing appreciation for district court feedback.⁵²⁷ An ALJ complained that the district court is an “anti-ALJ bench.”⁵²⁸ The hearing offices lack any sort of structured process for learning from district court opinions. One ALJ tracks district court decisions and forwards them to colleagues, but this ALJ told us that management has discouraged such initiatives.⁵²⁹ OGC provides training on leading causes of remands,⁵³⁰ although OGC-Hearing Office interaction is irregular.⁵³¹ ALJs and decision writers do get e-mails notifying them when the court of appeals issues relevant decisions.⁵³²

A similarity between the two hearing offices was ALJ dissatisfaction with the quality of claimant representation. ALJs in the high remand district⁵³³ and the low remand district⁵³⁴ gave these representatives mostly negative reviews.

2. Contrasts in the District Courts

The low and high remand districts contrast with each other in a number of interesting ways. Some have to do with how the two courts process their social security dockets. In the low remand district, magistrate judges handle almost all of the cases, and the caseload is immense – about sixty cases per year per magistrate judge.⁵³⁵ The result is a decision making structure that bears some similarities to agency adjudication. Several judges described the premium they place

most of the time she knows what she is supposed to do but just cannot do it, given time constraints. District judges do not understand these pressures. ALJ 20 at 3.

⁵²⁷ Only one ALJ we interviewed did not express frustration with district courts, and he had recently been appointed to his current position and lacked sufficient exposure to the district court to comment. ALJ 16 at 4.

⁵²⁸ ALJ 4 at 5; *see also* ALJ 5 at 5 (complaining that district judges have little respect for the agency process); ALJ 3 at 4 (describing the district as “extremely liberal”).

⁵²⁹ ALJ 3 at 4-5.

⁵³⁰ ALJ 5 at 5; ALJ 4 at 5; OGC Lawyer 5, Second Interview at 6.

⁵³¹ One OGC lawyer described OGC visits to hearing offices as “very infrequent” and “very few.” OGC Lawyer 5, Second Interview at 6.

⁵³² ALJ 16 at 4; ALJ 6 at 5; Decision Writer 2 at 4.

⁵³³ For negative reviews, *see* ALJ 5 at 3; ALJ 20 at 5; ALJ 3 at 2. For an equivocal review, *see* ALJ 6 at 2.

⁵³⁴ ALJ 11 at 2; ALJ 9 at 2.

⁵³⁵ One magistrate judge told us that the district “often” has the highest social security caseload in the United States. Federal Judge 19 at 1. The district ranks in the top ten for cases decided between 2010 and 2013, according to our data.

on expeditious case processing, just as ALJs do.⁵³⁶ The magistrate judges rely on permanent clerks, a professionalized staff with technical expertise quite different from inexperienced term clerks with their generalist legal education as a background.⁵³⁷ Magistrate judges also work through appeals themselves, without clerk involvement.⁵³⁸

The district court in the high remand district takes a different approach. It does not treat social security cases any differently than other civil actions, eschewing the sort of institutional response to this docket that the low remand district has taken to cope with high case loads. As a matter of longstanding custom,⁵³⁹ district judges decide these cases with no magistrate judge involvement at all. The docket works out to about twelve appeals per judge annually.⁵⁴⁰ No district judge with whom we spoke assigns a permanent clerk to social security cases, and most rejected the idea that the complexities of social security cases require clerks to have particular expertise.⁵⁴¹ Another salient difference is that the share of claimants represented is considerably lower in the high remand district—roughly 66%, by comparison to nearly all in the low remand district.

⁵³⁶ One magistrate judge described his two clerks as “rolling balls of butcher knives” for their ability to work through almost one-and-a-half cases each per week. Federal Judge 19 at 4. Another, a self-described “machine” for social security appeals, insisted that his job is to “move cases along as fast as possible.” Federal Judge 2 at 2, 4. A third told us that she spends 2-8 hours on a social security case, deciding from the bench after oral argument on every case. Federal Judge 1 at 3. This figure was by far the lowest that any federal judge reported. Another magistrate judge in the district reported spending 15-25 hours on a case. Federal Judge 19 at 4.

⁵³⁷ Federal Judge 19 at 4; Federal Judge 1 at 3. The district also has two full-time “social security clerks,” or staff attorneys who work on nothing but social security appeals. Federal Judge 19 at 4.

⁵³⁸ Federal Judge 19 at 4 (reporting that he has written “many, many opinions myself” without law clerk involvement); Federal Judge 1 at 3 (reporting that she splits appeals with her clerks); Federal Judge 2 at 2 (reporting that he handles all social security cases himself).

⁵³⁹ Federal Judge 4 at 4.

⁵⁴⁰ We calculated this rate with the assumption that senior judges do not take these cases.

⁵⁴¹ One district judge scoffed at the notion and described a clerkship devoted to social security appeals as a “gruesome job.” Federal Judge 15 at 3. Another district judge described a job that included a significant amount of social security work as “torturous.” Federal Judge 4 at 3; *see also* Federal Judge 18 at 2 (insisting that term clerks are capable of mastering social security law to the extent necessary); Federal Judge 17 at 1 (denying that permanent clerks are necessary and rejecting the idea that social security cases are complicated). One judge handles social security cases himself, however, and insisted to us that a recent law graduate has “no prayer” of handling appeals correctly. Federal Judge 10 at 3.

Federal judges in the low remand district for the most part gave the agency mixed reviews. One magistrate judge strongly criticized the quality of ALJ decisions,⁵⁴² but others insisted that “ALJs do a pretty good job.”⁵⁴³ Interestingly, neither the caseload nor middling impressions of decisional quality have sapped judicial tolerance for this litigation.⁵⁴⁴ “This area of law” is a “passion,” one magistrate judge said.⁵⁴⁵

Our interviews with judges in the high remand district revealed a different attitude. Two judges expressed no general impression, favorable or not, of the agency’s work.⁵⁴⁶ Others described what one judge called a culture of “significant distrust.”⁵⁴⁷ The district’s “operating principle” is that most cases will be remanded, one judge told us.⁵⁴⁸ “We’re generally suspicious of the agency,” another one said.⁵⁴⁹ A third complained that “the situation is a mess,”⁵⁵⁰ and a fourth insisted that ALJs have “very little credibility.”⁵⁵¹ One judge described a case coming out of one hearing office as an “automatic remand.”⁵⁵² Another insisted that “ALJs’ decisions almost always depart[] from governing legal standards in one way or another.”⁵⁵³

E. Reflections on District-Level Variations

Our study of district-level variations prompts several reflections. First and most obviously, the low and high remand districts lie in different circuits. This fact alone accounts for a good deal of their contrasting remand rates. Circuit boundaries, we discovered, have significant importance for remand rates.

⁵⁴² Federal Judge 19 at 3.

⁵⁴³ Federal Judge 2 at 1; Federal Judge 1 at 1.

⁵⁴⁴ Federal Judge 7 at 3.

⁵⁴⁵ Federal Judge 2 at 4. Other judges expressed similar views. Federal Judge 19 at 4; Federal Judge 1 at 4.

⁵⁴⁶ Federal Judge 17 at 1; Federal Judge 18 at 1.

⁵⁴⁷ Federal Judge 15 at 1.

⁵⁴⁸ Federal Judge 4 at 1.

⁵⁴⁹ Federal Judge 14 at 1.

⁵⁵⁰ Federal Judge 15 at 1.

⁵⁵¹ Federal Judge 10 at 1.

⁵⁵² Federal Judge 15 at 1.

⁵⁵³ Federal Judge 23 at 1.

Second, the two districts differ in their organization and deployment of judicial resources available for social security cases. The assembly line-like process of decision-making in the low remand district, staffed exclusively by magistrate judges and their permanent clerks, contrasts sharply with what happens in the high remand district, where district judges and their term clerks handle social security appeals no differently than the rest of the civil docket. Does the volume of social security cases in the low remand district inure magistrate judges and their long-time clerks to claimant suffering? Does the relative paucity of cases handled in the first instance by inexperienced clerks create unwarranted sympathy for claimants in the high remand district? Does volume give judges in the low remand district a more accurate sense of the makeup of the national disability claimant population, and thus a metric for evaluating claims that better approximates what an ALJ uses? Are judges in the high remand district misled by an unrepresentative selection of cases? Our quantitative results do not suggest that the answer to these questions is yes.

Third, our data strongly suggest that district-wide cultures or attitudes influence remand rates. Our quantitative findings as to the single outlier judge hypothesis indicate that judges tend to march in lockstep with others in the same district when it comes to deciding social security appeals. To put it differently, judges with particularly high remand rates tend to have district colleagues with similarly high remand rates. A simple explanation for district-level variations that stresses individual judicial idiosyncrasy is incorrect.

Other than circuit boundaries, what factors produce these district-wide cultures? Our intensive qualitative study of three districts, and the clear contrasts between the high and low remand districts, lead us to hypothesize that hearing offices generate decisions of varying quality for district courts to review. Common sense dictates that the health of a working environment

can have an important impact on the quality of work performed.⁵⁵⁴ If this is so, and if we did not by coincidence stumble upon the only hearing offices that contrast sharply,⁵⁵⁵ then the average strength of an appeal might differ from one hearing office to the next.⁵⁵⁶

Other work suggests that an agency's reputation is important to its success or lack thereof in the federal courts.⁵⁵⁷ Moreover, once a culture forms, it may be difficult to change, and it may exaggerate or amplify actual differences in the strength of appeals. Federal judges used to uneven ALJ decisions might well search all the more for errors that confirm their background impressions.⁵⁵⁸ As Richard Posner insisted in an opinion involving another high volume adjudication context, "[d]eference is earned; it is not a birthright."⁵⁵⁹ This sentiment lacks any formal basis in law, of course. The statutory section prescribing the substantial evidence standard of review does not peg the deference a court owes the agency to its evaluation of agency competence.⁵⁶⁰ But we expect that some of Judge Posner's colleagues harbor the same view, whether they acknowledge it consciously or not.⁵⁶¹ Conversely, a judge used to high-quality decisions may give a weaker one the benefit of the doubt. The weakness of decisions from a hearing office, in other words, could have an outsized influence on outcomes at a district court.

⁵⁵⁴ See, e.g., Malcolm Patterson, *Organizational Climate and Company Productivity: The Role of Employee Affect and Employee Level*, 77 J. OCCUPATIONAL & ORG. PSYCHOLOGY 193 (2004).

⁵⁵⁵ Our ODAR Region and labor market findings give us further reason to doubt such a coincidence.

⁵⁵⁶ Cases can get appealed to a district court from a variety of hearing offices, not just those located within the district's boundaries. Such remote decisions include those rendered by National Hearing Centers and those rendered by ALJs conducting hearings by video. We do not believe, however, that this fact poses a significant complication for our hypothesis. Even if, say, a third of appeals come from out-of-district hearing offices, no single hearing office will produce more appeals to a district court than those within its boundaries. The hearing offices generating the largest and most regular stream of appeals to a district court, we suggest, will most significantly influence the culture within that district.

⁵⁵⁷ HUME, *supra* note 273 at 117 ("Once an agency's reputation has deteriorated, it can be an uphill battle to restore its good name.").

⁵⁵⁸ On confirmation biases, see, e.g., DANIEL KAHNEMAN, THINKING, FAST AND SLOW 80-81 (2011).

⁵⁵⁹ *Kadia v. Gonzales*, 501 F.3d 817, 821 (7th Cir. 2007).

⁵⁶⁰ 42 U.S.C. § 405(g).

⁵⁶¹ For an example of a district judge who does share this attitude with respect to social security appeals, see *Hart v. Astrue*, Civ. No. 08-07, 2008 WL 3456864, at *10 (W.D. Wis. Aug. 11, 2008).

We recommend that the agency explore our hypothesis further. The agency could begin by determining, for example, whether patterns in Appeals Council decision making differ by hearing office, why they differ, and how they have changed over time.⁵⁶² Analysts at the Appeals Council are not instructed to review decisions from particular hearing offices or particular ALJs with more careful scrutiny.⁵⁶³ A number of claimant representatives and ALJs expressed the view to us that Appeals Council review is unpredictable and often seemingly arbitrary.⁵⁶⁴ If the Appeals Council catches errors from hearing offices at similar rates, this surface equivalence might mask inconsistent review that allows weak decisions from struggling hearing offices to enter the federal courts more often. The agency could also survey federal judges more extensively, to ask them for their perceptions of the quality of the ALJ decisions they review. If judicial satisfaction or dissatisfaction matches up with personnel or management issues in particular hearing offices, the correlation could indicate the uneven quality of pools of appeals.

Unless a federal judge is convinced that the social security docket he manages does not differ qualitatively from the docket his colleague in a different district handles, the federal judge is unlikely to find district-level inconsistencies troubling. If the inputs differ, the outputs should differ as well. Confirming or disproving our hypothesis is an important starting point to improve the consistency of district-court decision-making.

⁵⁶² The data available to us were insufficient for us to make this determination.

⁵⁶³ Agency Official 3, Third Interview at 2.

⁵⁶⁴ *E.g.*, Claimant Representative 11 at 2 (describing Appeals Council review as a “crapshoot”); Claimant Representative 2 at 3 (describing the value of Appeals Council review as “zip” and “random”); ALJ 2 at 2 (estimating that only 5-10% of Appeals Council remands express a valid concern with ALJ orders); ALJ 10 at 5 (insisting that Appeals Council review is not meaningful).

Part V. The Procedural Governance of Social Security Litigation

Even a cursory examination of social security litigation in the federal courts reveals a startling fact: the procedural governance of these cases is chaotic. District courts function as courts of appeals for social security claimants. The Federal Rules of Civil Procedure therefore do not work well for these cases, as they are designed for civil actions litigated in the first instance. As a result, districts and even individual judges have forged their own workarounds. A kaleidoscopic proliferation of procedures is the result.

Social security litigation involves claims governed by a national body of law and emerges from an administrative process that, at least formally, is the same everywhere. The lawyers who litigate these cases often have regional or even national practices, and they routinely appear in multiple districts. No good reason exists for procedural localism, and yet localism is the status quo. What results are inefficiencies, as lawyers constantly have to change how they litigate otherwise identical claims based on a district's or even an individual judge's particular procedural preferences. The procedural differences are often picayune – one district allows twenty page briefs, another eighteen – and thus involve little conceivable benefit. A social security practice is a high volume one. For lawyers on this assembly line, the costs of procedural differences, in terms of redundant effort and unintended errors, can mount quickly.

We provide illustrative examples of how the procedural governance of social security litigation varies in this Part. We discuss the costs of procedural localism and question whether any countervailing benefits justify them. A number of specific problems with particular practices came up in our interviews with OGC lawyers and claimant representatives. We report what we heard here. Ultimately this discussion supports our recommendation, provided in Part VI, that Congress legislate a set of procedural rules to provide these cases with the uniform national governance that they deserve.

None of our interview subjects suggested that procedural variation affects outcomes, with one now mostly obsolete exception.⁵⁶⁵ We therefore doubt that procedural variation explains intra-circuit differences among districts' remand rates. In addition, given the wide variety in orders even within districts, quantitative study of this topic would be difficult if not impossible. For these reasons, we have eschewed a quantitative analysis of the effects of different procedural rules. Statistics are unnecessary for our main point. The modern system of federal civil procedure is based in significant part on the premise that litigation proceeds most efficiently and successfully when procedural governance is uniform to the extent possible.⁵⁶⁶ Balkanized procedural governance of social security litigation conflicts with this fundamental premise.

A. Variations in Social Security Procedure

The Federal Rules of Civil Procedure provide a national and uniform set of procedural rules for general civil litigation in the federal courts. For certain types of specialized litigation, such as habeas corpus and bankruptcy, the Federal Rules provide only a baseline of procedural governance. Congress and other rule makers have crafted particularized, nationally applicable rules that address these cases' substance-specific issues.⁵⁶⁷ Other types of litigation also present specialized procedural needs, but rule makers have not seen fit to prescribe tailored procedural regimes for them.⁵⁶⁸

Social security litigation falls into the second category. These cases have a unique procedural posture. As appeals, they do not fit the Federal Rules' one-size-fits-all mold, crafted with an ordinary civil action as the model. No consensus exists, for instance, on exactly what

⁵⁶⁵ *E.g.*, Claimant Representative 5 at 3 (denying that an erstwhile rule in the WDNY requiring simultaneous briefing had an impact on outcomes); OGC Lawyer 12 at 3 (denying that differences in rules affect outcomes); OGC Lawyer 14 at 3 (same).

⁵⁶⁶ *E.g.*, Stephen B. Burbank, *Pleading and the Dilemmas of "General Rules"*, 2009 WIS. L. REV. 535, 536.

⁵⁶⁷ FED. R. CIV P. 81(a)(2) (noting that the Federal Rules apply in bankruptcy proceedings to the extent permitted by the bankruptcy rules); *id.* 81(a)(4) (same for habeas).

⁵⁶⁸ *See generally* Megan M. La Belle, *The Local Rules of Patent Procedure*, 47 ARIZ. ST. L.J. 63 (2015) (discussing the procedural needs that patent cases pose and arguing for the promulgation of a specialized set of rules for them).

sort of motion litigants should file to get a social security case adjudicated on the merits.⁵⁶⁹ In the Eastern District of Washington, the claimant files a motion for summary judgment,⁵⁷⁰ while a summary judgment motion is “not appropriate” in the Southern District of Iowa.⁵⁷¹

When specialized litigation with unique procedural needs lacks a tailored set of national rules for its governance, districts and even individual judges tend to step into the breach.⁵⁷² This is precisely what has happened with social security litigation. The Federal Rules exempt social security cases from the initial disclosures requirement of Rule 26,⁵⁷³ and they limit electronic access of nonparties to filings in these cases,⁵⁷⁴ but otherwise they treat disability cases with silence. A dizzying array of local rules, district-wide scheduling orders, and individual case management orders fills this procedural vacuum.⁵⁷⁵

These rules and orders address a multitude of issues at every stage in a case, as the following examples illustrate:

- The Northern District of California and the Eastern District of Wisconsin provide form complaints for plaintiffs to use in social security cases.⁵⁷⁶
- In the Western District of New York, the agency files the certified transcript of the administrative proceedings (“CAR”) in lieu of an answer within ninety days of the complaint’s service.⁵⁷⁷ In the Northern District of New York, the agency must file an answer along with the CAR within one hundred days of the complaint’s service.⁵⁷⁸

⁵⁶⁹ *E.g.*, Denlow, *supra* note 197 at 106-107 (providing examples of divergences among districts).

⁵⁷⁰ *E.g.*, Order Setting Schedule, *Donvan-Terris v. Colvin*, Civ. No. 14-5125, E.D. Wash., April 8, 2015. All case-specific orders referred to in this Report are on file with the authors.

⁵⁷¹ S.D. Iowa Local R. 56(i).

⁵⁷² *E.g.*, La Belle, *supra* note 568 at 86-92 (discussing the proliferation of local and individual rules for patent litigation).

⁵⁷³ FED. R. CIV. P. 26(a)(1)(B)(i).

⁵⁷⁴ FED. R. CIV. P. 5.2(c).

⁵⁷⁵ *E.g.*, Denlow, *supra* note 197 at 106-107.

⁵⁷⁶ Complaint for Judicial Review of Decision of Commissioner of Social Security, N.D. Cal.; Complaint for Review of a Final Decision by the Commissioner of the Social Security Administration, E.D. Wis.

⁵⁷⁷ Standing Order, In the Matter of Actions Seeking Review of the Commissioner of Social Security’s Final Decisions Denying Social Security Benefits, W.D.N.Y., Sept. 6, 2013, at 1.

⁵⁷⁸ General Order No. 18, In the Matter of: Northern District Order Directing Filing of Answer, Administrative Record, Briefs, and Providing for Oral Hearing on Appeal From Social Security Benefits Decision, N.D.N.Y., Sept. 12, 2003, at 1.

- Some judges in the Central District of California give the agency an automatic 30-day extension to file its response to the complaint.⁵⁷⁹ The Western District of New York gives the agency an automatic 30-day extension.⁵⁸⁰
- Most judges in the Central District of California require the plaintiff to send a “written and detailed proposal of settlement” to the agency after service of the administrative record. The parties have to meet and confer over this proposal before the agency files its answer.⁵⁸¹ Judges in the Eastern District of California have a similar requirement, although pursuant to a different timeline.⁵⁸² In the Southern District of California, social security cases are exempt from a local rule requiring early settlement conferences shortly after an answer is filed.⁵⁸³ Hence the parties proceed to merits briefing without a required attempt to settle.⁵⁸⁴
- In the Northern District of Iowa, the parties must file a joint stipulation of facts, “setting forth those facts found in the administrative record which are relevant to the issues presented” by the appeal, before filing their merits briefs. The claimant writes the first draft of this statement.⁵⁸⁵ The parties must file a “joint statement of facts” in the District of New Hampshire. The agency writes the first draft and circulates it after the claimant files her opening merits brief.⁵⁸⁶
- In the Eastern District of Missouri, the plaintiff files a “brief in support of the complaint” within thirty days of the answer’s service. The agency files its “brief in support of the answer” thirty days after service of the plaintiff’s brief. The plaintiff then has fourteen days to file a reply.⁵⁸⁷ In the Western District of Missouri, the equivalent periods are forty days, forty days, and twenty-one days.⁵⁸⁸
- Principal merits briefs in the Western District of Washington are limited to eighteen pages.⁵⁸⁹ Parties have thirty pages in the Eastern District of Wisconsin.⁵⁹⁰

⁵⁷⁹ *E.g.*, Case Management Order, _____ v. Astrue, Civ. No. 11-1819, C.D. Cal., Mar. 11, 2011, at 2.

⁵⁸⁰ Standing Order, In the Matter of Actions Seeking Review of the Commissioner of Social Security’s Final Decisions Denying Social Security Benefits, W.D.N.Y., Sept. 6, 2013, at 1.

⁵⁸¹ *E.g.*, Case Management Order, _____ v. Astrue, Civ. No. 11-396, C.D. Cal., Mar. 11, 2011, at 3.

⁵⁸² *E.g.*, Scheduling Order, Copeland v. Colvin, Civ. No. 15-96, E.D. Cal., Jan. 22, 2015, at 2 (requiring that the plaintiff file an “opening brief” if an exchange of “letter briefs” does not produce a settlement).

⁵⁸³ S.D. Cal. Local R. 16.1(e)(3).

⁵⁸⁴ *E.g.*, Order Setting Briefing Schedule, Worsham v. Colvin, Civ. No. 15-55, S.D. Cal., Apr. 9, 2015.

⁵⁸⁵ Briefing Schedule, N.D. Iowa, Oct. 15, 2015.

⁵⁸⁶ D.N.H. Local R. 9.1(c).

⁵⁸⁷ E.D. Mo. Local R. 56-9.02.

⁵⁸⁸ W.D. Mo. Local R. 9.1(e).

⁵⁸⁹ *E.g.*, Scheduling Order, Dodson v. Colvin, Civ. No. 15-5263, W.D. Wash. Oct. 16, 2015.

⁵⁹⁰ Procedures in Social Security Disability Appeals, United States District Court, Eastern District of Washington, Aug. 1, 2014, at 2.

- Some magistrate judges in the Central District of California require a “joint stipulation” in lieu of merits briefs. Scheduling orders for these judges prescribe in detail the content of these joint stipulations,⁵⁹¹ designed to ensure that the parties respond directly to each other’s contentions.⁵⁹² Other magistrate judges in the Central District of California require motions for summary judgment, not joint stipulations.⁵⁹³ Parties use summary judgment in other districts to litigate the merits of the claimant’s appeal.⁵⁹⁴ In still others, a motion for summary judgment is “not appropriate.”⁵⁹⁵ In a number of jurisdictions, parties litigate the merits of the claimant’s appeal with filings that more closely resemble what one would find in a court of appeals than anything provided for in the Federal Rules of Civil Procedure.⁵⁹⁶ In the District of New Hampshire, the merits brief is styled a “motion for order reversing decision of the Commissioner or for other relief.”⁵⁹⁷
- Some judges in the Southern District of Texas require the parties to submit cross motions for summary judgment simultaneously.⁵⁹⁸ Others require the claimant to move first and the agency to respond.⁵⁹⁹
- Until June 2015, the Eastern District of New York required the agency to file a motion for judgment on the pleadings before any merits filing by the claimant, requiring the agency effectively to guess at the problems that the claimant believed warranted the appeal. This rule remains in effect when claimants proceed pro se.⁶⁰⁰ A judge in the Western District of Virginia has the agency brief the case first and only requires a brief from the claimant if he or she waives oral argument.⁶⁰¹
- Some judges hold oral argument in every or almost every case.⁶⁰² In some districts, “generally oral argument will not be heard.”⁶⁰³

⁵⁹¹ *E.g.*, Case Management Order, _____ v. Astrue, Civ. No. 11-396, C.D. Cal. Mar. 3, 2011, at 4-7.

⁵⁹² Claimant Representative 8 at 5.

⁵⁹³ *E.g.*, Order, _____ v. Astrue, Civ. No. 11-1776, C.D. Cal., Mar. 9, 2011, at 2-3.

⁵⁹⁴ *E.g.*, Order Setting Schedule, Donvan-Terris v. Colvin, Civ. No. 14-5125, E.D. Wash., April 8, 2015.

⁵⁹⁵ S.D. Iowa Local R. 56(i); *see also* E.D. Mo. Local R. 56-9.02.

⁵⁹⁶ *E.g.*, Scheduling Order, Walker v. Colvin, Civ. No. 15-5252, W.D. Wash., Sept. 9, 2015, at 2-3; General Order No. 18, In the Matter of: Northern District Order Directing Filing of Answer, Administrative Record, Briefs, and Providing for Oral Hearing on Appeal From Social Security Benefits Decision, N.D.N.Y., Sept. 12, 2003, at 1-2; E.D. Mo. Local R. 56-9.02; D.N.H. Local R. 9.1

⁵⁹⁷ D.N.H. Local R. 9.1

⁵⁹⁸ *E.g.*, Order, McCord v. Colvin, Civ. No. 14-208, S.D. Tex., June 9, 2014.

⁵⁹⁹ *E.g.*, Briefing Schedule, Barnes v. Colvin, Civ. No. 14-482, S.D. Tex., Sept. 3, 2014, at 1-2.

⁶⁰⁰ Administrative Order 2015-05, In re: Scheduling in Social Security Cases, E.D.N.Y., June 19, 2015.

⁶⁰¹ Standing Order Governing Development of Social Security Cases Assigned to Judge Conrad, W.D. Va., Jan. 1, 2005.

⁶⁰² Federal Judge 1 at 3.

⁶⁰³ General Order No. 18, In the Matter of: Northern District Order Directing Filing of Answer, Administrative Record, Briefs, and Providing for Oral Hearing on Appeal From Social Security Benefits Decision, N.D.N.Y., Sept. 12, 2003, at 3; *see also* W.D. Mo. Local R. 9.1(e) (noting that the court will decide a case “without oral argument, unless otherwise directed”); Order, _____ v. Astrue, Civ. No. 11-1776, C.D. Cal., Mar. 9, 2011, at 3 (“The Court will take the motions under submission without oral argument, unless the Court otherwise orders.”).

B. Problems With Disuniformity

The most prominent of commentators in civil procedure have long criticized procedural governance by local rules, general orders, standing orders of individual judges, scheduling orders for particular cases, and so forth.⁶⁰⁴ A complete rehearsal of their many objections is unnecessary here, but several find particular purchase for social security litigation. Local rules and the like create “legal clutter,” or “background noise” of procedural variation that consumes time and distracts lawyers from focusing on the merits.⁶⁰⁵ A lawyer practicing in different districts or before different judges within the same district may constantly have to toggle back and forth among various requirements for the same task. As one OGC lawyer told us, this inefficient juggling detracts from the quality of representation.⁶⁰⁶ “Difficult local rules,” she insisted, are a “huge waste of time.”⁶⁰⁷ Until recent changes improved the situation, the two dozen magistrate judges in the Central District of California used something like seventeen

⁶⁰⁴ E.g., Comment, *The Local Rules of Civil Procedure in the Federal District Courts – A Survey*, 1966 DUKE L.J. 1011, 1012 & n.6 (quoting a letter from Charles Alan Wright referring to local rules as the “‘soft underbelly’ of federal procedure”); Charles Alan Wright, Foreword, *The Malaise of Federal Rulemaking*, 14 REV. LITIG. 1, 11 (1994) (referring to the proliferation of local rules among the districts and lamenting that that “procedural anarchy is now the order of the day”); Geoffrey Hazard, *Undemocratic Legislation*, 87 YALE L.J. 1284, 1285 (1978) (chiding local rules as little more than “measurements of the chancellor’s feet”); John P. Frank, *Local Rules*, 137 U. PA. L. REV. 2059, 2060 (1989) (denouncing variations among local rules as an “outright abomination” and blasting the “sheer arrogance and irresponsibility” of judges for fashioning them).

⁶⁰⁵ Paul D. Carrington, *A New Confederacy? Disunionism in the Federal Courts*, 45 DUKE L.J. 929, 947 (1996).

⁶⁰⁶ OGC Lawyer 12 at 3. Cf. OGC Lawyer 8 at 4 (insisting that consistency within a district is “very helpful”). The great litigator John Frank described “a litigation law office” with “an entire wall plastered with the separate local rules for each of the district courts in [its] community.” Frank, *Local Rules*, *supra* note 604 at 2060. OGC lawyers described a version of this phenomenon to aid their efforts at litigating in the Central District of California. See also Claimant Representative 8 at 2 (complaining that it can be “tedious” to try to follow the many case management orders in the Central District of California).

⁶⁰⁷ OGC Lawyer 12 at 5.

different case management orders.⁶⁰⁸ Under such conditions, OGC “can’t zealously represent its client,” a lawyer lamented.⁶⁰⁹

This procedural clutter is more than a mere annoyance, as an example illustrates. Cases routinely raise the same issues, so rather than reinvent the wheel, lawyers often borrow text from one brief for another. By testing arguments in a number of cases and redeploying them, lawyers can raise the quality of their representation. Moreover, these practices lower the cost of social security litigation, both for claimants and the government. OGC lawyers rarely spend more than three days writing a merits brief, and claimant representatives must work on a tight budget. These shortcuts save invaluable time. But procedural variation makes economizing difficult, forcing lawyers to refashion even a successful argument to fit a judge’s procedural parameters. Not only does such redundant drafting take time, it also may compromise the quality of advocacy. And all to a questionable end, so that a brief will not exceed eighteen pages instead of twenty, or so that an argument fits in a joint submission instead of an ordinary merits brief.

Lawyers cannot adapt to these variations by reducing the number of courts in which they litigate. To suggest that they do so ignores the institutional organization of the social security bar. Some claimant representatives litigate exclusively in one or two districts, particularly those in big cities that attract a large volume of litigation.⁶¹⁰ But others, particularly practitioners who specialize in federal court litigation,⁶¹¹ need to litigate in multiple districts to earn a living. By making their practices economically viable, multi-jurisdictional litigation also probably creates access to qualified counsel for claimants in smaller locales. Each of OGC’s regions handles a

⁶⁰⁸ OGC Lawyer 15 & OGC Lawyer 25 at 2. *See also* Claimant Representative 8 at 2 (complaining that it can be “tedious” to try to follow the many case management orders in the Central District of California).

⁶⁰⁹ OGC Lawyer 21 at 2; OGC Lawyer 26 at 1 (complaining that different scheduling orders “trip” people up); *see also* Carl Tobias, *Charles Alan Wright and the Fragmentation of Federal Practice and Procedure*, 19 *YALE L. & POL’Y REV.* 463, 465 (2001).

⁶¹⁰ *E.g.*, Claimant Representative 3 at 1; Claimant Representative 4 at 1.

⁶¹¹ *E.g.*, Claimant Representative 9 at 1; Claimant Representative 10 at 1.

number of districts. OGC lawyers rarely practice exclusively in only one district. Their workloads are a function not of what might be desirable, but of variables beyond their control, such as the number of cases appealed and resources available to the agency.

Another problem with localism is inconsistency, or the possibility that like cases will not be treated alike.⁶¹² In only one instance did one of our interview subjects suggest that an idiosyncratic practice, the EDNY's erstwhile rule requiring the agency to brief a case first, impacted outcomes.⁶¹³ But inconsistencies beyond outcome variation should concern the courts. Differences among rules and practices create arbitrary delay for certain claimants. Some judges grant extensions readily, for example, while others are sticklers for deadlines.⁶¹⁴ An OGC lawyer might have briefs due on the same day, one for a stickler judge and another for a permissive judge. This lawyer naturally will prioritize the brief for the stickler judge, knowing that he can always get an extension from the permissive judge.⁶¹⁵ The second case's claimant has to wait longer than the first for her claim's resolution, having already pursued her claim for two years in the agency. The requirement some districts have that the plaintiff request a voluntary remand by letter before writing the merits brief prompts a similar worry. If these "pre-motion letters" force OGC lawyers to dig into a case earlier than they might otherwise,⁶¹⁶ the letters might prompt earlier RVRs. A claimant litigating before a judge in the Eastern District of

⁶¹² Robert E. Keeton, *The Function of Local Rules and the Tension with Uniformity*, 50 U. PITT. L. REV. 853, 874 (1989).

⁶¹³ OGC Lawyer 5, Second Interview at 1.

⁶¹⁴ *E.g.*, Briefing Schedule, *Davis v. Colvin*, Civ. No. 15-1023, N.D. Iowa, Oct. 8, 2015 ("No extensions will be granted without exceptional cause."); OGC Lawyer 10 at 4 (observing that a magistrate judge who handles a large volume of social security cases is reluctant to grant extensions); OGC Lawyer 11 at 4 (observing that some judges are "sticklers" for extensions and that others grant them freely); OGC Lawyer 27 at 5 (describing some judges who never give extensions and others who do so freely).

⁶¹⁵ OGC Lawyer 27 at 6.

⁶¹⁶ OGC lawyers disagreed on whether these letters accomplish this goal. *Compare* OGC Lawyer 2 at 4 (insisting that these letters are a waste of time), *with* OGC Lawyer 15 & OGC Lawyer 25 at 6 (conceding that sometimes the early review these letters require prompt an earlier settlement). *See also* Claimant Representative 9 at 5 (suggesting that these letters have no effect).

California who requires these letters may get a voluntary remand sooner than an identical claimant in the Eastern District of Washington, which does not.

Finally, procedural localism is problematic because courts and judges might generate requirements without sufficient deliberation or opportunities for public input.⁶¹⁷ A proposed amendment to the Federal Rules of Civil Procedure must proceed through several steps until its promulgation. It requires the support of multiple constituencies, and the process affords the public several opportunities to comment. Local rules need to pass muster “in the eyes of only one set of beholders,” the local rules committee for the district.⁶¹⁸ Even worse are “general orders” for a district, an umbrella term that describes district-wide orders issued by some process less formal than local rulemaking. Local rules at least require some public comment period before promulgation.⁶¹⁹ A district can issue a general order without any process at all. The Northern District of Iowa, for example, issued a generic “Briefing Schedule” on October 15, 2015, that requires the parties to file a joint stipulation of facts.⁶²⁰ The district issued it without an opportunity for public input. Had it invited reactions from social security litigators, the district surely would have discovered universal antipathy for such a procedure.⁶²¹

Worst of all is the delegation of procedural governance to case management or scheduling orders issued by individual judges. A judge can change her guidelines overnight, without any warning, and without even the acquiescence of or input from her judicial

⁶¹⁷ The opacity of procedural governance for social security litigation is reflected in a mundane but telling phenomenon. We had significant difficulty even locating many of the procedural requirements districts or individual judges have set for social security litigants. General orders containing these requirements are often available on district websites, but we found them only with significant, time-intensive searching. Some districts have agreed on a generic scheduling order, but we could only find it by downloading it from individual docket sheets on PACER. In one instance, a district advertised on its website that it had just promulgated a new set of requirements for social security cases. This notice did not include a link to the order itself, which we could only obtain by calling the clerk’s office and asking for a copy by e-mail.

⁶¹⁸ Hazard, *supra* note 604 at 1286.

⁶¹⁹ FED. R. CIV. P. 83; *Hollingsworth v. Perry*, 558 U.S. 183, 190 (2010).

⁶²⁰ Briefing Schedule, _____ v. Colvin, N.D. Iowa, Oct. 15, 2015 (on file with authors).

⁶²¹ See *infra* at text accompanying notes 624-640

colleagues.⁶²² One magistrate judge in the Central District of California changed scheduling orders four times within a single year, we were told, each time without notice or warning to lawyers.⁶²³ Judges often do not even bother posting these orders on their webpages. This problem of notice is endemic. Good luck to the lawyer who wants to assemble a comprehensive set of procedural rules for the judges in her district or circuit. This is tough sledding.

C. Problems With Particular Procedures

We would hesitate to recommend reforms were lawyers generally content with the procedural governance of social security litigation, however balkanized it is. But interview subjects consistently expressed frustration. Concerns about the following issues surfaced repeatedly.

1. Joint Submissions

Joint submissions include both joint statements of facts⁶²⁴ and joint stipulations.⁶²⁵ Parties typically submit the former in addition to merits briefs, to memorialize the factual background to an appeal in a single document authored by both sides.⁶²⁶ The latter replace dueling merits briefs. The parties write and file a single document containing both of their arguments.

Several rationales ostensibly explain these departures from ordinary adversarial procedure. Judges in the Central District of California tired of the two sides failing to engage directly with each other's arguments.⁶²⁷ To remedy this problem, about half of the magistrate judges in the district require the parties to go back-and-forth in the same document. These

⁶²² OGC Lawyer 15 & OGC Lawyer 25 at 2.

⁶²³ OGC Lawyer 15 & OGC Lawyer 25 at 2.

⁶²⁴ *E.g.*, Briefing Schedule, _____ v. Colvin, N.D. Iowa, Oct. 15, 2015; Briefing Schedule and Format Order, Dillon v. Colvin, Civ. No. 15-5034, D.S.D., July 10, 2015.

⁶²⁵ *E.g.*, Case Management Order, _____ v. Astrue, Civ. No. 11-424, C.D. Cal., Mar. 15, 2011, at 4.

⁶²⁶ *E.g.*, D.N.H. Local R. 9.1(b)-(e).

⁶²⁷ OGC Lawyer 28 at 6; Claimant Representative 8 at 5.

judges' orders go so far as to dictate a format that generates a sort of point-counterpoint engagement.⁶²⁸ Presumably districts and judges require joint statements of facts for the same reason. Also, as one claimant representative told us, the requirement that the parties confer on a joint statement of facts before merits briefing, or at least before the agency files its brief, is thought to produce earlier RVRs.⁶²⁹

Perhaps no practice attracted as much criticism in our interviews as these joint submissions.⁶³⁰ Even the sole claimant representative who expressed support for the requirement admitted that colleagues in the plaintiff's bar either dislike it, or they saddle the agency with inconvenience and burden by failing to follow rules for these documents.⁶³¹ These submissions, which an OGC lawyer described as neither "neither joint nor stipulated,"⁶³² are problematic for several reasons. First, they require a "remarkably cumbersome process."⁶³³ The parties must coordinate on a number of issues, including how many pages each side gets,⁶³⁴ who files what exhibits, and even which merits arguments are raised.⁶³⁵ In a high volume litigation context, this mandated cooperation adds a layer of inefficiency that busy attorneys can ill afford.

⁶²⁸ A typical order requires the following format:

- I. Issue No. 1
 - A. Plaintiff's Contentions Regarding Issue No. 1
 - B. The Commissioner's Contentions Regarding Issue No. 1
 - C. Plaintiff's Reply Regarding Issue No. 1.

See, e.g., Case Management Order, _____ v. Astrue, Civ. No. 11-424, C.D. Cal., Mar. 15, 2011, at 7.

⁶²⁹ Claimant Representative 10 at 2.

⁶³⁰ *E.g.*, OGC Lawyer 14 at 3; OGC Lawyer 9 at 4; OGC Lawyer 7 at 3; Claimant Representative 10 at 2; OGC Lawyer 12 at 3.

⁶³¹ Claimant Representative 8 at 5.

⁶³² OGC Lawyer 29 at 5.

⁶³³ OGC Lawyer 15 & OGC Lawyer 25 at 1.

⁶³⁴ An OGC lawyer complained to us of claimants writing 30-page drafts in jurisdictions where the documents cannot exceed 40 pages, thus leaving OGC only 10 pages. OGC Lawyer 14 at 3. Another lawyer described bloated, 80-page drafts arriving in their inboxes, effectively requiring them to join a document that will all but surely anger the judge. OGC Lawyer 15 & OGC Lawyer 25 at 7.

⁶³⁵ An OGC lawyer complained that the joint submission model makes it difficult affirmatively to raise dispositive issues that plaintiffs understandably do not mention themselves. OGC Lawyer 15 & OGC Lawyer 25 at 3. *Cf.* Claimant Representative 8 at 5 (conceding that the agency faces this problem).

A claimant representative described joint submissions as “royal waste of time,” and she “can’t stand working on them.”⁶³⁶

Second, the joint submission process effectively outsources the unpleasant job of enforcing deadlines to the lawyers.⁶³⁷ A scheduling order might specify the dates on which each version of the draft might be exchanged, but also provide that nothing should be filed until the final submission is ready. A lawyer who misses a deadline is supposed to request an extension, to allow the entire briefing schedule to reset. But he sometimes fails to do so, a lapse surely encouraged at least in part by the fact that the judge remains in the dark until the submission finally gets filed. The other side has to hound her adversary to request the extension, lest her time to respond get shortened.⁶³⁸ One can readily imagine the reluctance of a lawyer to bring the deadline issue to the court’s attention, worried that doing so will irritate a judge impatient with such tedious housekeeping matters.

The joint submission process may indeed force the parties to engage more directly with each other’s argument. There is no evidence, however, that they prompt more or earlier RVRs.⁶³⁹ Moreover, other vehicles, namely pre-motion letters, can accomplish this objective more efficiently.

District courts require joint submissions of various sorts in other litigation settings. But rarely do the parties actually litigate the merits of their case through a joint submission. A number of districts require the parties to submit statements of fact at summary judgment, to identify those that are disputed and undisputed. These rules have a certain logic; summary judgment requires the identification of undisputed facts from a morass of discovery materials.

⁶³⁶ Claimant Representative 10 at 2.

⁶³⁷ OGC Lawyer 28 at 6.

⁶³⁸ OGC Lawyer 15 & OGC Lawyer 25 at 3.

⁶³⁹ OGC Lawyer 29 at 5; Claimant Representative 10 at 2.

But social security litigation involves a closed, already created record. It is revealing that no court of appeals finds something like a joint statement of facts necessary to navigate a record.⁶⁴⁰

2. Simultaneous Briefing

Simultaneous briefing is another practice that drew criticism from OGC lawyers and claimant representatives alike.⁶⁴¹ Some judges in the Southern District of Texas, for example, require both sides to submit cross motions for summary judgment. They then respond to each other's briefs in simultaneously filed responses.⁶⁴² This practice effectively doubles the number of briefs the lawyers must file. OGC lawyers find themselves in a particularly strange position when they draft their summary judgment motions, because they have to anticipate the issues the claimant believes merit an appeal. The result, a "bland" brief, as one OGC lawyer described it,⁶⁴³ surely adds little value on top of what the brief filed in response to the claimant's motion creates. When a defendant in an ordinary civil action moves for summary judgment, it knows from the complaint what claims the plaintiff asserts. In social security cases, complaints are almost invariably form documents with little specificity. OGC lawyers have to guess what problems the claimant will identify.

Until recently, the Eastern District of New York had an even more curious practice, requiring that the agency file the opening brief and claimants respond. A judge in Virginia continues to employ this "affirmative briefing" process.⁶⁴⁴ OGC lawyers hate affirmative

⁶⁴⁰ The Federal Civil Rules Advisory Committee abandoned a proposal to require such statements in Rule 56, based in part on concerns about burden and cost. Edward H. Cooper, *Revising Rule 56: Judge Mark R. Kravitz and the Rules Enabling Act*, 18 LEWIS & CLARK L. REV. 591, 603-604 (2014).

⁶⁴¹ OGC Lawyer 13 at 3; OGC Lawyer 9 at 4; OGC Lawyer 1 at 3.

⁶⁴² *E.g.*, Order, *McCord v. Colvin*, Civ. No. 14-208, S.D. Tex., June 9, 2014.

⁶⁴³ OGC Lawyer 13 at 3.

⁶⁴⁴ Standing Order Governing Development of Social Security Cases Assigned to Judge Conrad, W.D. Va., Jan. 1, 2005.

briefing, and they believe it actually affects outcomes.⁶⁴⁵ The requirement makes little sense. An OGC lawyer has to write a comprehensive opening brief, lest he irritate the court, while not spoon-feeding issues the plaintiff's lawyer might otherwise overlook.⁶⁴⁶ An OGC lawyer told us that he dealt with this situation by trying to say as little as possible in his opening brief.⁶⁴⁷ This result is obviously undesirable.

3. Pre-Motion Letters

A number of jurisdictions require the parties to exchange pre-motion letters, or letters written before full-dress merits briefing, to present each side's case in a concise manner.⁶⁴⁸ These letters require the parties to delve into the case, albeit in an abbreviated way, earlier than they might otherwise. Faster RVRs might result, or so the logic goes.

Pre-motion letter requirements have drawn mixed reviews. Some lawyers think they are a waste of time and do not prompt settlement,⁶⁴⁹ while others agree that they sometimes yield quicker RVRs.⁶⁵⁰ An OGC lawyer complained that a pre-motion letter requirement forces her to digest a lengthy CAR twice, once to answer the claimant's letter and again to write the merits brief. She works on so many cases between the two reviews that the earlier one does not really streamline the latter.⁶⁵¹ A claimant representative, in contrast, insisted that the time he spends on

⁶⁴⁵ OGC Lawyer 5, Second Interview at 1. *Cf.* OGC Lawyer 6 at 1 (describing the EDNY as a claimant-friendly jurisdiction because of the affirmative briefing requirement). Affirmative briefing has earned judicial criticism as well. *Hamilton v. Secretary of Health & Human Servs.*, 961 F.2d 1495, 1501 (10th Cir. 1992); Federal Judge 10 at 1 (explaining that an affirmative briefing rule "never made sense"). A claimant representative we interviewed does not see the reason for it. Claimant Representative 4 at 4.

⁶⁴⁶ OGC Lawyer 4 at 2.

⁶⁴⁷ OGC Lawyer 4 at 2.

⁶⁴⁸ *Eg.*, Scheduling Order, *Copeland v. Colvin*, Civ. No. 15-96, E.D. Cal., Jan. 22, 2015, at 2; *see also* Claimant Representative 10 at 2; OGC Lawyer 2 at 4.

⁶⁴⁹ OGC Lawyer 2 at 4; Claimant Representative 9 at 5.

⁶⁵⁰ Claimant Representative 10 at 2.

⁶⁵¹ OGC Lawyer 28 at 7. Other OGC lawyer conceded that sometimes the first review can streamline the second one. OGC Lawyer 15 & OGC Lawyer 25 at 6.

the letters is not wasted, even if an RVR doesn't result, because it reduces the time he then spends on the merits brief.⁶⁵²

4. Oral Argument

Most judges do not require or hold oral argument in social security cases, but the practice may be gaining in popularity.⁶⁵³ A magistrate judge from the Northern District of Mississippi has advocated for oral argument, coupled with decisions from the bench, as a way to work through a large docket quickly.⁶⁵⁴ Claimant representatives did not express any particular dislike of the practice to us.⁶⁵⁵ OGC lawyers, in contrast, almost uniformly dislike oral argument.⁶⁵⁶ Again, the problem involves duplication of effort. Having worked up a case months earlier when drafting the merits brief, the lawyer must retrace her steps to bone up before oral argument.⁶⁵⁷ Judges often conduct oral argument by phone, but sometimes attorneys must appear in person. In such instances, travel can create an additional inconvenience, especially given that OGC lawyers cover multiple districts from a single regional office.⁶⁵⁸

5. Deadlines, Page Limits, and Extensions

Briefing schedules and rules about page limits and extensions diverge considerably. A judge in the Southern District of Texas issued an order giving the claimant four months from the date of the CAR's filing to file his opening brief; the government then had twenty-eight days to respond.⁶⁵⁹ The standard scheduling order in the Western District of Washington gives both

⁶⁵² Claimant Representative 1 at 4.

⁶⁵³ Federal Judge 1 at 2-3 (noting that she has begun holding oral arguments and deciding from the bench over the past year).

⁶⁵⁴ This judge has created videos to train colleagues in this technique. A copy of the video is on file with the authors.

⁶⁵⁵ Claimant Representative 5 at 3; Claimant Representative 6 at 3; Claimant Representative 9 at 5.

⁶⁵⁶ *E.g.*, OGC Lawyer 13 at 3; OGC Lawyer 1 at 2; OGC Lawyer 7 at 3; OGC Lawyer 8 at 3.

⁶⁵⁷ OGC Lawyer 13 at 3; OGC Lawyer 7 at 3; OGC Lawyer 8 at 3.

⁶⁵⁸ OGC Lawyer 13 at 3.

⁶⁵⁹ Briefing Schedule, *Barnes v. Colvin*, Civ. No. 14-482, S.D. Tex., Sept. 3, 2014.

parties thirty days,⁶⁶⁰ as does an order used in the Central District of California.⁶⁶¹ In the Eastern District of Wisconsin and the Western District of Missouri, the claimant has forty days after the CAR's filing to file her brief, and the agency then has forty days within which to respond.⁶⁶² The EDNY and WDNY give both sides sixty days,⁶⁶³ while the NDNY gives them forty-five days.⁶⁶⁴ As for page limits, they range all over the map. A merits brief cannot exceed eighteen pages in the Western District of Washington, for example, while the EDNY allows for unlimited pages.⁶⁶⁵

We have alluded to the problem with variations in extension practices above. As for deadlines, they “cr[y] out for standardization,” an OGC lawyer told us.⁶⁶⁶ Given the high caseloads that social security lawyers handle, the proliferation of idiosyncratic briefing schedules “adds a layer of administrative complexity,” or the sort of procedural clutter that localism engenders.⁶⁶⁷ The same is true of page limits. A lawyer who litigates in multiple districts constantly must recalibrate how extensively she can argue a particular issue. An argument developed for a district with more permissive limits, for example, will have to be retooled for a district with stricter ones.

⁶⁶⁰ *E.g.*, Scheduling Order, *Dodson v. Colvin*, Civ. No. 15-5263, W.D. Wash., Oct. 16, 2015, at 1.

⁶⁶¹ Order, _____ v. *Astrue*, Civ. No. 11-1776, C.D. Cal., Mar. 9, 2011, at 3.

⁶⁶² Procedures in Social Security Disability Appeals, ED. Wis., Aug. 1, 2014; W.D. Mo. Local R. 9.1(e).

⁶⁶³ Administrative Order 2015-05, In re: Scheduling in Social Security Cases, E.D.N.Y., June 19, 2015; Standing Order, In the Matter of Actions Seeking Review of the Commissioner of Social Security's Final Decisions Denying Social Security Benefits, W.D.N.Y., Sept. 5, 2013.

⁶⁶⁴ General Order No. 18, In the Matter of: Northern District Order Directing Filing of Answer, Administrative Record, Briefs, and Providing for Oral Hearing on Appeal From Social Security Benefits Decision, N.D.N.Y., Sept. 12, 2003, at 1-2.

⁶⁶⁵ *Compare* Scheduling Order, *Dodson v. Colvin*, Civ. No. 15-5263, W.D. Wash., Oct. 16, 2015, at 1, *with* Federal Judge 4 at 5.

⁶⁶⁶ OGC Lawyer 5, First Interview at 5.

⁶⁶⁷ OGC Lawyer 1 at 5.

Part VI. Recommendations and Suggestions

This Report's charge directed us to study social security litigation in the federal courts. Our recommendations and suggestions reflect this focus. They address aspects of social security litigation that could be rendered more efficient, and they target the agency's relationship with the federal courts. Surely these proposed reforms will strike some as weak medicine for what ails the social security disability system. Only a very small number of claimants who initially apply for benefits ultimately seek judicial review. To most, a uniform set of procedural rules for social security cases will have little significance. Likewise, the agency struggles with vast numbers of claims and a mandate to implement a national policy consistently. Improved communication with the federal courts will not relieve these challenges. Still, nearly 20,000 social security appeals proceed each year in the federal courts. This litigation can change for the better.

Our suggested reforms divide into two categories. By our understanding, ACUS's recommendations typically involve suggested legislation or rules, not changes to internal agency governance. We thus limit our recommendations, offered for ACUS's consideration, to improvements to the procedural rules governing social security litigation, to changes to how the federal courts and the agency communicate, and to adjustments to OGC's role. Our research nonetheless lead us to believe that the agency could improve its standing in the federal courts, to claimants' betterment, with adjustments to its internal processes. We offer these changes as informal suggestions that the SSA should consider as it continues with its efforts to improve the disability claims adjudication process.

A. Recommendations for ACUS Consideration

1. Independent Litigating Authority

Recommendation 1. Congress should give the Social Security Administration independent litigating authority.

The SSA should be given independent litigating authority, which it presently lacks.⁶⁶⁸ It cannot represent itself in the federal courts and must work through the Department of Justice.⁶⁶⁹ This limitation has two significant consequences. The SSA cannot decide on its own when to appeal an adverse district court decision, and the SSA must enter an appearance through a U.S. Attorney's office.

To the extent that the agency is dissatisfied with the overall tenor of social security case law, its negligible access to appellate review must shoulder some of the blame. This lack of access may have significantly blunted the force that the substantial evidence standard of review might otherwise seem to have, at least by its terms. A district court's characterization of something as legal error is unlikely to be appealed. Also, the asymmetric threat claimants can make to appeal may steer district and magistrate judges to favor claimants in close cases.

The influence circuits have over district court remand rates places the agency's lack of appellate access in particularly stark relief. Claimants set the entirety of each circuit's social security agenda. If one, such as the Seventh Circuit, tilts in a pro-claimant direction, a docket constructed by claimant representatives will only encourage it to tilt even further so. The results are ever-rising remand rates for the districts within the circuit.

We take no position on any particular treatment of social security law in existing precedent. More agency appeals could have several ameliorative effects, even if one were indifferent to the overall tenor of social security case law.⁶⁷⁰ First, disuniformity in case law complicates the agency's capacity to maintain and implement a consistent national policy. It also means differential treatment of otherwise identical claimants that depends upon accident of

⁶⁶⁸ For a list from 2012 of agencies with independent litigating authority, see DAVID E. LEWIS & JENNIFER L. SELIN, ADMINISTRATIVE CONFERENCE OF THE UNITED STATES: SOURCEBOOK OF UNITED STATES EXECUTIVE AGENCIES 116 n.296 (DEC. 2012)

⁶⁶⁹ *Id.* at A-26.

⁶⁷⁰ If the agency is right to believe that a particular body of case law tilts in a pro-claimant direction, more appeals are an imperative.

geography. More regular appellate review, whether secured by claimants or the agency, should help. “[T]he importance of maintaining harmony among the Circuits on issues of law” is an important consideration when a circuit addresses an issue for the first time.⁶⁷¹ As a corollary, arguments about the need for national uniformity likely will gain more traction at the circuit level than at the district level, where judges’ first and often only obligation is to their circuit’s precedent.

Second, appellate review might force the agency to clarify or reconsider its own interpretation of the Act or regulations, or to respond to adverse decisions with rulemaking. District court decisions rarely, if ever, prompt such reconsideration.⁶⁷² A circuit decision does so. The decision may trigger the agency’s acquiescence process, for one thing; its effects will be felt within the agency more informally as well. Although ALJs and decision writers are not supposed to follow case law unless the agency says so, those we interviewed were familiar with circuit decisions. Any single district court decision is too insignificant, except in a rare instance, to prompt such a reaction.

The traditional justifications for DOJ control over agency appeals are less convincing in the social security context.⁶⁷³ The federal government may have a legitimate concern that it speak with one voice when it litigates issues that cut across subject areas, such as those involving Congress’s Commerce Clause power or the appropriate pleading standard under Rule 8 of the

⁶⁷¹ *Wong v. PartyGaming Ltd.*, 589 F.3d 821, 827-28 (6th Cir. 2009); *see also Wheeler v. Pilgrim’s Pride Corp.*, 591 F.3d 355, 363 (5th Cir. 2009) (observing in an en banc decision reversing an earlier panel’s explicit decision to create a circuit split that the court should “look to the opinions of other circuits for persuasive guidance, always chary to create a circuit split”); *United States v. Auginash*, 266 F.3d 781, 784 (8th Cir. 2001); *Andrews v. Chevy Chase Bank*, 545 F.3d 570, 576 (7th Cir. 2008); *American Vantage Companies, Inc. v. Table Mountain Rancheria*, 292 F.3d 1091, 1098 (9th Cir. 2002).

⁶⁷² *Cf. Soc. Sec. Ruling 96-1p* (indicating that the agency will essentially ignore district court decisions inconsistent with its policy unless the court directs otherwise in a class action).

⁶⁷³ For a general discussion of the traditional justifications, see Neal Devins & Michael Herz, *The Uneasy Case for Department of Justice Control of Federal Litigation*, 5 U. PA. J. CONSTIT. L. 558, 570-94 (2003); *id.* at 594 (observing that these justifications “seem overblown”).

Federal Rules of Civil Procedure. But social security appeals rarely involve such issues and almost invariably focus instead on detailed and technical questions of social security law and policy.⁶⁷⁴ Also, a fear that an agency unbound might squander the federal government's appellate capital by crying wolf too often is simplistic. An entity as sophisticated as a court of appeals can differentiate among types of federal government litigants.

From time to time a social security appeal might involve an issue whose resolution could affect the federal government's legal standing in other contexts. The agency might argue, for example, that an interpretation of the Act issued in a particular kind of document deserves deference. The resolution could affect how courts within the circuit address such "*Chevron* Step Zero" issues for other agencies' interpretations. To ensure that the agency appreciates the broader governmental perspective, the SSA could be required to give the DOJ notice of an appeal sufficient to enable the DOJ to intervene when necessary.

The other concern with the SSA's lack of independent litigating authority involves OGC lawyers' obligation to work through U.S. Attorney's Offices. In our view, this partnership creates inefficiencies and produces uncertain benefits. The involvement of skilled AUSAs may improve briefs in some instances, but many OGC lawyers whom we interviewed, while appreciative of the help, conceded that AUSAs rarely contribute much in the way of substance. The U.S. Attorney may advocate for the agency on other issues, such as local rule changes, and AUSAs' familiarity and relationships with individual judges surely benefits lawyers, like those in OGC, who have to cover a wider territory. These benefits, however, are insufficient to warrant the arrangement that currently persists, especially in light of our other

⁶⁷⁴ *Id.* at 604 (arguing that, in such circumstances, the argument in favor of independent litigating authority in the agencies is strongest).

recommendations.⁶⁷⁵ Among other problems, we detected confusion among federal judges in our interviews over who exactly does what on behalf of the government in social security cases. Congress should allow the SSA to enter its own appearances and relieve U.S. Attorneys of the unnecessary, redundant, and sometimes confusing role they currently shoulder.

An individual U.S. Attorney's office can mostly remove itself from disability litigation without Congressional action by designating OGC lawyers as SAUSAs. This work-around is imperfect, if better than the status quo in districts that do not use SAUSAs. The U.S. Attorney's name still appears on a brief when an OGC lawyer appears as a SAUSA. An office's reluctance to surrender all control over what gets filed is understandable. The office may fear that its reputation could suffer if OGC lawyers file poor-quality briefs. The best alternative is to remove U.S. Attorney involvement altogether. Courts would have no reason to attribute shortcomings in OGC work product to the DOJ or the U.S. Attorney's Office. Independent litigating authority would clarify that OGC, not the U.S. Attorney, bears responsibility for litigation decisions and for the quality of government briefing, and thus is the proper target for any judicial irritation.

If Congress will not vest the SSA with independent litigating authority, then an imperfect but helpful substitute is the universal designation of OGC lawyers as SAUSAs in all districts. A U.S. Attorney's Office could profitably shift lawyer resources toward other, less redundant ends. For the most part,⁶⁷⁶ the federal judges we interviewed expressed satisfaction with the quality of agency lawyering. Briefs so deficient as to threaten the U.S. Attorney's reputation are probably infrequent, if our interviews with federal judges are any indication. If OGC lawyering does raise concerns with federal judges, the U.S. Attorney could always withdraw the SAUSA designation.

⁶⁷⁵ If the Administrative Office of the U.S. Courts and the agency collaborate on the creation of district standing committees, for example, these committees would lessen the need for the U.S. Attorney's assistance in local district matters, such as drafting local rules for social security cases.

⁶⁷⁶ One federal judge described OGC lawyering as "poor," complaining that OGC lawyers rely too much on boilerplate in their briefs and do not address medical facts at a sufficient level of specificity. Federal Judge 13 at 1-2.

2. Social Security Rules Enabling Legislation

Recommendation 2. Congress should enact enabling legislation to clarify the U.S. Supreme Court's authority to promulgate procedural rules for social security litigation. The Judicial Conference should authorize the appointment of a social security rules advisory committee, and the U.S. Supreme Court should approve a set of social security rules drafted by this committee.

The proliferation of inconsistent procedural rules for social security litigation results from a dearth of national procedural governance. Disability appeals are just that – appeals. The Federal Rules of Civil Procedure are designed for civil actions of first impression, and the processes they contemplate work poorly for cases better suited to appellate rules.

The case for a single, national set of rules for social security litigation is strong. The substance of these cases differs very little from one part of the country to the next. They emerge from a single, national administrative process, one that produces the same sort of record for review everywhere. Procedural needs rarely vary from one social security case to the next. There are a lot of these cases. During the twelve months that ended on September 30, 2014, petitioners filed 19,185 “general” habeas corpus petitions in the federal courts.⁶⁷⁷ This large chunk of the district court docket poses special procedural needs, and therefore a particular set of rules sensibly governs it.⁶⁷⁸ During the same time period, social security claimants filed 19,146 appeals in district courts.⁶⁷⁹ In addition, the lawyers who litigate these cases often have a regional or even national practice. This institutional fact is certainly true of OGC, and it is often true of claimant representatives as well.

Finally, any marginal return from local experimentation with procedural governance for social security cases has long since disappeared. Districts have tinkered for long enough. At this

⁶⁷⁷ Table C-2A, U.S. District Courts – Civil Cases Commenced, by Nature of the Suit, During the 12-Month Periods Ending September 30, 2009 Through 2014.

⁶⁷⁸ E.g., Rules Governing Section 2254 Cases in the United States District Courts, at www.uscourts.gov.

⁶⁷⁹ Table C-2A, U.S. District Courts – Civil Cases Commenced, by Nature of the Suit, During the 12-Month Periods Ending September 30, 2009 Through 2014.

point, the only conceivable beneficiaries of procedural localism are judges with idiosyncratic preferences, and they might be mistaken about the benefits their requirements ostensibly create. Localism creates inefficiency and detracts from the lawyer resources available for merits briefing, and these judges may actually be making their own jobs harder.

As discussed below, we recommend that a uniform set of procedural rules replace the chaos of local rules, standing orders, and individual practices that presently govern social security litigation in the federal courts. More important than any particular rule, however, is the existence of an ongoing rulemaking process that can create and revise social security rules as procedural needs arise and change. The Rules Enabling Act delegates power to craft rules of practice and procedure to the Supreme Court. This power is then exercised through procedures crafted by the Judicial Conference of the United States.⁶⁸⁰ Congress should amend the Rules Enabling Act to clarify that this delegation of rulemaking power includes the power to craft social security rules. A committee authorized by the Judicial Conference would then shoulder an ongoing responsibility for these rules.

Five options exist for the promulgation of a uniform set of procedural rules for social security litigation. First, the agency and/or a national claimant representative organization could market an agreed-upon set of best practices to individual judges, for adoption in case management orders or the like. This option is least promising. A judge-by-judge effort will almost certainly leave balkanized procedures in place, and the problem of procedural localism will persist.

A second option is a district-by-district approach. A district could adopt these recommendations in a general order or as amendments to its local rules. Local rulemaking is

⁶⁸⁰ 28 U.S.C. §§ 2072(a), 2073(a). The discussion that follows simplifies the rulemaking process to a certain extent. For its details, see, e.g., <http://www.uscourts.gov/rules-policies/about-rulemaking-process/how-rulemaking-process-works>.

preferable because it requires districts to solicit public input and cannot proceed entirely behind closed doors.⁶⁸¹ Also, amendments to local rules require more process than general orders, and for this reason may prove more resistant to judicial whim. Finally, an individual judge will have less leeway to vary a practice when it is prescribed in a local rule instead of a standing order or the like. In the absence of either of amendments to the Federal Rules of Civil Procedure or legislation, the broad adoption of a set of model local rules is the best option.

A district-by-district effort at local rules amendment would prove laborious, and surely some districts might resist. A third option is for Congress simply to legislate the procedural rules applicable in social security cases. There is some precedent for this sort of lawmaking.⁶⁸² It is an imperfect choice, however, because legislation without an ongoing committee responsible for amendments as the need arises would freeze the rules in time. If a rule proved problematic, or if a new rule were needed, Congress would have to pass another statute.

A fourth option is to amend the Federal Rules of Civil Procedure to include specialized provisions for social security cases where appropriate. The Rules Enabling Act prohibits a rule that would “abridge, enlarge, or modify any substantive right.”⁶⁸³ This limitation has disabled rule makers from crafting a heightened pleading standard for certain categories of claims, for example.⁶⁸⁴ The committees responsible for rules of procedure and practice, however, have appropriately recognized that certain areas of litigation simply have different procedural needs that do not implicate the contours of substantive rights. They have responded with appropriately tailored rules that account for these differences without privileging plaintiffs or defendants in any

⁶⁸¹ FED. R. CIV. P. 83(a)(1).

⁶⁸² Congress legislated a set of habeas corpus rules in 1976, although their origin and its relationship with the ordinary rulemaking process is complicated. *See generally* Robert N. Clinton, *Rule 9 of the Federal Habeas Corpus Rules: A Case Study on the Need for Reform of the Rules Enabling Acts*, 63 IOWA L. REV. 15 (1977).

⁶⁸³ 28 U.S.C. § 2042(b).

⁶⁸⁴ David Marcus, *The Past, Present, and Future of Trans-Substantivity in Federal Civil Procedure*, 59 DEPAUL L. REV. 371, 414 (2010).

systematic fashion.⁶⁸⁵ Indeed, local rules are subject to Enabling Act limitations.⁶⁸⁶ The many districts with social security-specific local rules have either gravely erred in their judgments that such rules are permissible, or the Enabling Act’s limits permit what we recommend here.

Nonetheless, we recommend a fifth option for a mundane but not unimportant reason. There is no obvious place to locate particularized social security rules in the Federal Rules of Civil Procedure, in part because they do not include provisions designed for the equivalent of appellate practice. Many existing rules would have to be amended, in ways that would change their character fundamentally. Rule 3, for example, presently reads, “a civil action is commenced by filing a complaint with the court.”⁶⁸⁷ It has been amended only once since 1938, to adjust its style.⁶⁸⁸ To account for the recommendations proposed in this Report, Rule 3 would have to be revised to say something like, “Except in actions brought under 42 U.S.C. § 405(g), a civil action is commenced by filing a complaint with the court.” Adjustments to Rules 8, 12, 56, and others would be necessary as well. Rather than fill the Federal Rules with such clutter, the Supreme Court could approve an amendment to Rule 81(a) adding social security appeals to a list of actions to which the Federal Rules apply only to the extent consistent with a specialized procedural regime.

A social security rules advisory committee supervised by the Judicial Conference, with members appointed by the Chief Justice, should craft this specialized procedural regime. The benefits of such rulemaking are legion. They include the capacity to draw on empirical research to inform rule amendments; a process that invites robust participation from stakeholders; and a commitment to reasoned, expert deliberation. The Supreme Court might already have the

⁶⁸⁵ See, e.g., FED. R. CIV. PROC. 23.1; FED. R. APP. PROC. 13; *id.* 15.1.

⁶⁸⁶ 28 U.S.C. § 2071(a)-(b); 12 CHARLES ALAN WRIGHT ET AL., FEDERAL PRACTICE & PROCEDURE § 3153 (updated 2016).

⁶⁸⁷ FED. R. CIV. PROC. 3.

⁶⁸⁸ FED. R. CIV. PROC. 3, Adv. Comm. Note.

requisite rulemaking power under the Rules Enabling Act, and likewise the Judicial Conference might already have sufficient authority to create such a committee.⁶⁸⁹ To ensure that the Enabling Act's "substantive right" limitation does not deter the Supreme Court or the Judicial Conference, however, Congress should amend the act to clarify that its delegation of rulemaking power includes the power to craft social security rules. Legislation enacted in 1964 to give the Court the power to craft bankruptcy rules offers a model.⁶⁹⁰ The Judicial Conference could then authorize the appointment of a social security rules advisory committee, and the Supreme Court could eventually approve a uniform set of rules for this litigation.

3. Social Security Rules

Recommendation 3. A uniform set of procedural rules for social security litigation should contain (a) a rule requiring the claimant to file a notice of appeal instead of a complaint; (b) a rule requiring the agency to file the certified administrative record instead of an answer; (c) a rule requiring the parties to exchange merits briefs instead of motions; (d) a rule setting appropriate deadlines and page limits; and (e) a rule creating a presumption against oral argument.

Once Congress has acted and a social security rules advisory committee is in place, this committee and the other institutions involved in rulemaking should consider the following proposed rules for the uniform procedural governance of social security litigation.

a. Complaints and Answers

Claimants should not have to file complaints in social security cases, and the agency should not have to file answers. For the most part these documents are needless formalities.⁶⁹¹

⁶⁸⁹ The Supreme Court has suggested that the exercise of rulemaking power to craft a specialized set of rules for a particular area of litigation may be appropriate under the Rules Enabling Act. *Harris v. Nelson*, 394 U.S. 286, 300 n.7 (1969) (inviting the Federal Civil Rules Advisory Committee to draft a set of procedural rules for habeas corpus litigation).

⁶⁹⁰ 28 U.S.C. § 2075 ("The Supreme Court shall have the power to prescribe by general rules, the forms of process, writs, pleadings, and motions, and the practice and procedure in cases under title 11."); *see also* Alan N. Resnick, *The Bankruptcy Rulemaking Process*, 70 AM. BANKR. L.J. 245 (1996); Lawrence P. King, *The History and Development of the Bankruptcy Rules*, 70 AM. BANKR. L.J. 217 (1996).

⁶⁹¹ *E.g.*, Claimant Representative 6 at 3 (observing that the answer never serves a function). One claimant representative noted that, if the agency has an affirmative defense like *res judicata*, he needs to know of its existence before spending twenty-five hours writing a merits brief. Claimant Representative 1 at 5. This issue could be

Both sides file boilerplate documents that serve no purpose other than to satisfy the Federal Rules' requirements.⁶⁹² The exercise of preparing an answer would make sense if it forced OGC lawyers to conduct an early defensibility analysis and thereby prompted earlier RVRs in appropriate cases. But none of our interview subjects believes that this outcome materializes.⁶⁹³ The pleadings are "useless," to quote a federal judge,⁶⁹⁴ and no claimant representative reported to us that he or she used information in answers to prepare his or her briefs.

Claimants should commence social security cases the same way an appeal begins in the circuits, with the filing of a single-page notice of appeal.⁶⁹⁵ The agency gleans no useful information from a boilerplate complaint, so this convenience would prove costless. In lieu of an answer, the agency should simply file the certified administrative record ("CAR"),⁶⁹⁶ a practice that a number of districts already allow,⁶⁹⁷ and one that resembles what the Federal Rules of Appellate Procedure contemplate for other agency appeals.⁶⁹⁸ This alternative to the ordinary process for a case's commencement would save the lawyers unnecessary work.⁶⁹⁹

b. Merits Briefs

Districts require very different procedural vehicles for litigating the merits of the claimant's appeal. The most-used vehicle, summary judgment,⁷⁰⁰ suits social security litigation

resolved with an appropriate rule requiring the agency to identify any affirmative defenses when it produces the CAR.

⁶⁹² OGC Lawyer 1 at 5 (boilerplate); OGC Lawyer 12 at 2 (observing that judges never rely on anything but the merits briefs); OGC Lawyer 4 at 2 (insisting that the process is meaningless); Claimant Representative 1 at 5 (acknowledging that answers are "pretty formulaic").

⁶⁹³ OGC Lawyer 9 at 2 (exchanging complaints and answers does "not help[] us figure out the truth and reach justice").

⁶⁹⁴ Federal Judge 9 at 4.

⁶⁹⁵ FED. R. APP. PROC. 3(a); *id.* Form 1 (providing an illustrative notice of appeal from a district court decision); *id.* Form 3 (providing an illustrative petition for review from an agency decision).

⁶⁹⁶ *Cf.* OGC Lawyer 9 at 2 (insisting that it "would be fantastic" to be able to file the CAR in lieu of an answer);

⁶⁹⁷ *E.g.*, General Order 05-15 Re: Social Security Cases, Actions Seeking Review of the Commissioner of Social Security's Final Decision Denying An Application for Benefits, W.D. Wash., June 1, 2015.

⁶⁹⁸ FED. R. APP. PROC. 17.

⁶⁹⁹ OGC Lawyer 4 at 2; OGC Lawyer 6 at 2; OGC Lawyer 9 at 2.

⁷⁰⁰ Denlow, *supra* note 197 at 106.

poorly. A magistrate judge aptly summarized Rule 56's poor fit. Summary judgment asks whether undisputed facts require the entry of judgment as a matter of law, whereas social security appeals require judges to review contested factual findings, however deferentially. Evidence unearthed in discovery gets appended to summary judgment motions; a social security case requires a judge to review a closed record. A judge can deny a summary judgment motion and let a case proceed to trial, whereas a social security appeal requires a decision on the merits.⁷⁰¹

Appellate procedure provides an obvious alternative.⁷⁰² The parties should simply exchange and file merits briefs, a simple process for a simple procedural task. The district court functions as an appellate tribunal. The parties base their arguments on a closed record, and the district court employs appellate standards of review for questions of law and fact. Parties do not engage in discovery. A case that requires a more elaborate procedure, such as what Rule 56 provides, is rare to the point of vanishing.

A vehicle for merits litigation modeled on appellate advocacy may trouble those judges who prefer either cross motions for summary judgment or joint submissions. The rationale for a cross motions requirement, which simply doubles the amount of briefing the parties must do, eludes us. Courts may have fashioned joint submission requirements for worthy reasons. There are no data to suggest, however, that joint submissions have facilitated earlier requests for voluntary remand. We appreciate that a joint submission forces the agency to address the claimant's arguments point-by-point. An appellate briefing approach would not accomplish the same end as mechanically. But judges can always hold lawyers responsible for disorganized briefing. If an agency brief fails to address a claimant's argument, or vice versa, the judge can

⁷⁰¹ *Id.* at 106.

⁷⁰² *Id.* at 126.

treat the failure as a concession. A few such penalties would likely incentivize lawyers to take more care with their briefs. Moreover, courts of appeals surely wade through disorganized and nonresponsive opposition briefs all the time, and they cannot respond by requiring a joint submission instead of ordinary merits briefs.⁷⁰³ Judges who single out social security litigation for this particularly inefficient and burdensome procedure visit costs on these lawyers that others do not have to bear.

c. Deadlines and Page Limits

The Federal Rules of Appellate Procedure require the appellant to file its opening brief within forty days of the record's filing. The appellee then has thirty days to file its response.⁷⁰⁴ While circuit briefs will often involve more complicated issues than a typical social security case, this schedule offers helpful guidance. The typical appellate lawyer probably has lower caseloads than claimant representatives or OGC lawyers. Also, there is something unfair in allowing claimant representatives more time to brief than OGC lawyers enjoy. We thus believe that a modest expansion from the 40-day baseline is appropriate. Claimants should have sixty days after service of the CAR to file their opening briefs, and OGC lawyers should have sixty days to respond.

The Federal Rules of Appellate Procedure limit principal briefs to 14,000 words.⁷⁰⁵ This length is excessive for social security appeals, which tend to involve the same legal issues repeatedly. Judges do not require an elaborate introduction to the law,⁷⁰⁶ especially if the agency

⁷⁰³ Rule 32(e) of the Federal Rules of Appellate Procedure requires courts of appeals to accept briefs that comply with the form requirements of the rule. A circuit rule requiring joint submissions in lieu of briefs would be inconsistent with Rule 32(e).

⁷⁰⁴ FED. R. APP. PROC. 31(a)(1).

⁷⁰⁵ FED. R. APP. PROC. 32(a)(7).

⁷⁰⁶ Federal Judge 12 at 4 (explaining that experienced clerks do not need five pages of boilerplate to be educated about social security law); Federal Judge 20 at 6 (insisting that lawyers spend too much time providing information that judges do not need); Federal Judge 11 at 2 (explaining that page limits helped cut down on briefs with material that judges do not need); *id.* at 2 (explaining that judges do not need boilerplate about the five steps); Federal Judge

and claimant representatives draft a guide to social security law as we suggest below, boilerplate discussions of background law and the standard of review are unnecessary. With some exceptions, OGC lawyers and claimant representatives alike expressed comfort with a limit of 20-25 pages, or about 5,000-6,250 words.⁷⁰⁷ Every judge we interviewed, save one,⁷⁰⁸ indicated that 25 pages is reasonable but long enough.⁷⁰⁹ Indeed, several judges suggested that they view briefs any longer than this limit as evidence of subpar social security lawyering.⁷¹⁰ We thus recommend a page limit of 6,250 words, or about 25 pages of double-spaced text in Times New Roman 12-point font.

d. Oral Argument

16 at 3 (insisting that lawyers include too much in their briefs); Federal Judge 6 at 4 (insisting that “everyone knows” the basic law). Lawyers should take note of these and other comments that federal judges made to us. One claimant representative insisted that he had to include basic information in order to orient an inexperienced clerk to the nuts-and-bolts of social security law. Claimant Representative 7 at 4. If our interviews are any guide, this intuition is probably mistaken, and the extra text may simply create resentment, not goodwill.

⁷⁰⁷ Claimant Representative 4 at 4-5 (indicating satisfaction with 25 pages); Claimant Representative 2 at 4 (indicating satisfaction with fifteen pages); Claimant Representative 10 at 2 (indicating satisfaction with 25 pages); OGC Lawyer 11 at 4 (indicating satisfaction with 25 pages but suggesting that the limit could be lower); OGC Lawyer 26 at 1 (suggesting that a 20-page limit is fine but that 25 pages are too many); OGC Lawyer 10 at 4 (expressing satisfaction with 20 pages); OGC Lawyer 2 at 3 (suggesting that 20 pages is fine but 25 pages is too long). *But see* Claimant Representative 7 at 3 (describing a 25-page limit as “nuts”). *Cf.* Claimant Representative 1 at 4 (expressing an understanding for why briefs should be limited to 25 pages but worrying that in occasional cases with a large number of alleged errors the limit will create difficulties).

⁷⁰⁸ One judge expressed frustration with his district’s rule limiting briefs to 25 pages. To his mind, the rule addressed a nonexistent problem and doesn’t reflect anything but “animus” toward these cases. Federal Judge 3 at 3.

⁷⁰⁹ Federal Judge 13 at 3 (25 pages reasonable); Federal Judge 12 at 4 (25 pages reasonable); Federal Judge 20 at 6 (25 pages reasonable); Federal Judge 6 at 3 (25 pages reasonable); Federal Judge 7 at 4 (25 pages reasonable); Federal Judge 16 at 3 (20 pages reasonable); Federal Judge 19 at 5 (25 page limit not unreasonable, but most lawyers can do the job in many fewer pages); Federal Judge 4 at 5 (25 pages); Federal Judge 17 at 1 (20 pages).

⁷¹⁰ Federal Judge 12 at 4 (suggesting that only “lousy lawyers” submit a 40-page brief); Federal Judge 20 at 6 (suggesting that an author impugns her own writing ability if she cannot write a brief of less than 25 pages); Federal Judge 19 at 5 (“It’s better to operate with a rifle than a shotgun.”); Federal Judge 17 at 1 (insisting that if a lawyer can’t make an argument in twenty pages, the lawyer is “in trouble”).

Oral argument is a tricky issue. OGC lawyers do not like it. The claimant representatives we interviewed are comfortable with it. Whether to hold oral argument is mostly, if not always,⁷¹¹ a matter of individual judicial prerogative.

Oral argument imposes three types of costs. First, judges who hold in-person oral arguments saddle costs, in terms of travel and time, on lawyers who often litigate in multiple jurisdictions.⁷¹² Second, an oral argument might proceed months after the parties file their merits briefs. The lawyers have to relearn their cases, a task that is not insubstantial given the significant volume of business social security lawyers handle. Third and most important are the substantial costs that judges who decide cases on the bench after oral argument create. We appreciate that this method may enable a judge to work through her social security docket more quickly, as it obviates the need for a written order. But ALJs have more difficulty understanding bench decisions than written ones. OGC and the Appeals Council have to spend more effort to translate bench decisions into instructions to ALJs on remand, potentially creating even more delay for the claimant. Also, the agency must order a transcript, adding to the delay and expense. We therefore recommend a rule providing that the court will decide an appeal without oral argument unless specific and compelling reasons exist for why oral argument in a particular case would provide assistance to the court.

e. Additional Procedural Issues

Our interview subjects mentioned several other procedural issues of concern. One involves variations in extension practices. A claimant may have to wait longer for a decision than someone with similar alleged impairments, just because his case happens to be assigned to a

⁷¹¹ *E.g.*, D.N.H. Local Rule 7.1(d) (“Except as otherwise provided, the court shall decide motions without oral argument. The court may allow oral argument after consideration of a written statement by a party outlining the reasons why oral argument may provide assistance to the court.”).

⁷¹² These costs are less when arguments proceed by telephone or videoconference. If a rule discouraging oral argument is not adopted, one requiring courts to permit appearances by telephone would be helpful. We understand that videoconferencing can sometimes prove technically cumbersome for the agency.

judge who freely grants extensions. Lawyers with heavy caseloads naturally turn to these cases last. Also, OGC lawyers may be less willing to request voluntary remands when cases proceed before judges who either refuse to grant extensions or who require substantial advance notice before issuing them. If a judge requires seven days' notice, and if an OGC lawyer doesn't delve into a case in detail until a few days before a briefing deadline, the OGC lawyer may feel obliged to brief the case even if she might otherwise seek an RVR.⁷¹³ OGC lawyers are reluctant to justify a last-minute extension request on grounds that the ALJ's decision seems indefensible, and that they would prefer a voluntary remand. If the Appeals Council denies the RVR, the claimant and the court will know that the OGC lawyer may not fully believe in her own arguments.⁷¹⁴

We appreciate the need for flexibility with regard to extension requests and thus do not propose a rule regulating them. Nonetheless, we suggest that judges permit at least one extension as of right, without requiring parties to request extensions more than a day in advance. Expectations that lawyers ask four or five days before briefing deadlines mesh with the RVR process poorly and may actually depress the number of voluntary remands. Moreover, lawyers are understandably anxious about asking for an extension at the last minute.⁷¹⁵ Beyond this first request, a rule requiring more advance notice for additional extensions strikes us as sensible. A judge who freely allows three or even four extensions saddles claimants who have already waited for a decision for years with even more delay.⁷¹⁶

⁷¹³ OGC Lawyer 11 at 4. One OGC lawyer who litigates before a judge who requires 10 days' notice for extensions says she simply doesn't seek them. She defends even cases that otherwise might be voluntarily remanded "the best I can." OGC Lawyer 10 at 1.

⁷¹⁴ OGC Lawyer 10 at 4.

⁷¹⁵ Cf. Claimant Representative 4 at 5 (observing that it is a "bad practice" to request last-minute extensions).

⁷¹⁶ Cf. Claimant Representative 1 at 4 (insisting that he dislikes requesting extensions because his clients are "suffering").

A second issue involves pre-motion letters. A number of factors drive a district's RVR rate, including the propensity of OGC lawyers in that region to request voluntary remands. It stands to reason, however, that a mechanism that prompts lawyers to engage with the merits of a claim earlier will produce earlier, and thus more efficiently procured, settlements. In districts that lack these requirements, OGC lawyers often begin their review of a case only days before a briefing deadline.⁷¹⁷ If an OGC lawyer decides to RVR the case at this point, she does so months into litigation and after the claimant representative has spent time on and billed hours to a merits brief.⁷¹⁸

A couple of OGC lawyers conceded that a pre-motion letter can produce expedited RVRs “if taken seriously.” A two-page letter from the claimant, with issues presented in bullet point format, can work. The problem, they continued, is that plaintiffs often either send boilerplate letters, bereft of citations to the record and any arguments tailored to particular cases, or write lengthy, thirty page letters that differ in no material respect from the merits briefs themselves.⁷¹⁹ We also appreciate that the majority of OGC lawyers dislike pre-motion letter requirements.

At this point more study is needed to determine if pre-motion letters are worth the additional time and effort that they require. A social security rules advisory committee could examine this issue by comparing RVR practices and surveying lawyers in districts that do and do not require them.

4. Education and Communication

Recommendation 4. The Administrative Office of the United States Courts, the Federal Judicial Center, the Administrative Conference of the United States, and the Social Security Administration should cooperate on several initiatives to improve

⁷¹⁷ E.g., OGC Lawyer 2 at 2 (reporting that he does a defensibility analysis “usually a few days ahead” of a brief’s due date); OGC Lawyer 26 at 1-2 (reporting that, because of her workload, she has been “a week ahead” in her work only once during her multiple years working for OGC); OGC Lawyer 10 at 1 (observing that she begins her work on cases “just a few days out” and suggests her colleagues do as well).

⁷¹⁸ Cf. Claimant Representative 12 at 6 (suggesting an early settlement conference idea for these reasons).

⁷¹⁹ OGC Lawyer 15 & OGC Lawyer 25 at 5; see also OGC Lawyer 28 at 6.

communication among the agency, claimant representatives, and the judiciary, and to educate the judiciary in important aspects of the claims adjudication process. These initiatives should include the creation of social security standing committees for each district and the drafting of an introductory manual on social security law and processes.

Judicial ignorance of agency policymaking and adjudicative processes within the agency might contribute to cultures of distrust that exaggerate the remand rate. Not long ago, a judge in the Eastern District of Wisconsin complained of “far too many administrative law judges who are not conversant in Seventh Circuit law” when he chided the U.S. Attorney for attempting to defend certain ALJ decisions. This judge indicated little understanding of the agency’s policy toward circuit decisions that it believes to conflict with the Act and regulations.⁷²⁰ ALJs get reversed all the time but ignore us, another federal judge bemoaned to us, suggesting a similar lack of understanding.⁷²¹ A number of judges deprecated boilerplate in ALJ decisions,⁷²² without an apparent acknowledgement of the agency’s attempt to use the FIT template to improve decision-making and make it more consistent nationally.⁷²³

Better communication could have several helpful effects. First, the agency might respond more favorably to an adverse court decision if agency officials and lawyers believed that the courts are knowledgeable and respectful of their processes and the demands they shoulder.⁷²⁴ Federal courts sometimes temper their application of trans-substantive administrative law doctrines, such as *Chevron* or the substantial evidence standard of review, based on detailed knowledge of how a particular agency’s processes work.⁷²⁵ The same might be so here. In some instances, more knowledge might improve the agency’s standing in one judge’s eyes; in

⁷²⁰ Transcript, Mar. 13, 2013, at 14, at www.ssaconnect.com/260-sanctions.

⁷²¹ Federal Judge 20 at 6.

⁷²² *E.g.*, Federal Judge 10 at 1; Federal Judge 24 at 3; Federal Judge 8 at 2; Federal Judge 12 at 1.

⁷²³ OGC Lawyer 30 at 8 (complaining about judicial complaints that lack appreciation for what the FIT template is).

⁷²⁴ HUME, *supra* note 273 at 46 (reporting results of interviews with agency officials and observing “that when judges appear interested in their work and are knowledgeable about their procedures, administrators are more likely to accept decisions”).

⁷²⁵ See Richard H. Pildes, *Institutional Formalism and Realism in Constitutional and Public Law*, 2013 SUP. CT. REV. 1, 21-29.

another's, it might diminish.⁷²⁶ Either way, to the extent that judges calibrate their review based on knowledge of agency processes or a lack thereof, more education and communication can only help.

Each district should create a standing committee, whose membership could include at least one district or magistrate judge, an OGC lawyer who litigates regularly in the district, a respected claimant representative, and personnel from each hearing office within the district. This committee would ensure a regular and useful flow of information. Federal judges could explain their concerns about recurring issues, such as the use of problematic boilerplate.⁷²⁷ Agency lawyers could explain why the boilerplate remains or what steps have been taken to improve it.⁷²⁸ If a federal judge has questions about record development, she could ask them of the agency personnel and the claimant representatives. Claimant representatives and OGC lawyers could get feedback on briefing and procedural concerns. OGC lawyers could explain the RVR process to judges and thereby make sense of last minute extension requests that may otherwise prove confounding to judges.⁷²⁹ All could discuss what local rules and standing orders work and which ones cause problems.

⁷²⁶ Federal Judge 8 at 3 (expressing disappointed surprise when we described how decision writing works in hearing offices).

⁷²⁷ For decisions issued over five years criticizing the same boilerplate that apparently hinges a plaintiff's symptom evaluation in part on the ALJ's RFC finding, see, e.g., *Treichler v. Comm'r of Social Security*, 775 F.3d 1090, 1102-03 (9th Cir. 2014); *Parker v. Astrue*, 597 F.3d 920, 924 (7th Cir. 2010); *Hertz v. Colvin*, Civ. No. 13-6449, 2015 WL 5773722, at *3-4 (W.D.N.Y. Sept. 30, 2015); *Bales v. Colvin*, Civ. No. 14-297, 2015 WL 5690763, at *6 (E.D. Okla. Sept. 28, 2015); *Kaighn v. Colvin*, 13 F. Supp. 3d 1161, 1173-74 (D. Colo. 2014); *Suess v. Colvin*, 945 F. Supp. 2d 920, 928 (N.D. Ill. 2013); *Little v. Colvin*, Civ. No. 12-300, 2013 WL 2489173, at *5-6 (E.D. Va. June 7, 2013).

⁷²⁸ An OGC lawyer told us that the circular credibility analysis that so many courts have criticized has been addressed, and that decisions are now "not supposed to be written that way." OGC Lawyer 18 at 9. The committee would be a good place for OGC to communicate this and explain when judges could expect to see this language disappear in the decisions they review.

⁷²⁹ An OGC lawyer told us of an episode involving a magistrate judge who tired of last minute extension requests. The judge adopted a rule that required extension requests to be filed five days in advance of a briefing deadline. A "management team" had to meet with the magistrate judge, to explain the realities of agency processes. OGC Lawyer 2 at 2.

The only existing channels for communication among various players in the disability appeals process are either opinions or informal episodes, such as conversations at conferences or bench-bar meetings. Neither is effective. The latter are too irregular, and they do not necessarily include all relevant stakeholders. The former are problematic too. ALJs who dismiss the value of judicial feedback might be particularly inclined to ignore opinions they find harsh, disrespectful, or otherwise unduly critical. An opinion is also a one-way street and affords ALJs no opportunity to respond. Nor does it allow claimant representatives to voice concerns about problems in hearing offices for which district court remands are weak medicine. Also, the federal courts generate thousands of decisions each year, and the agency may have difficulty identifying from all of these the more urgent problems judges have identified. Rather than wait until the decisions accumulate – a period of years, perhaps, during which judicial frustration may deepen and harden – judges can urge reforms sooner. Lastly, without a different channel for information flow, efforts in the agency to improve in response to judicial criticism will remain opaque to judges for a couple of years, due to the lag time between an ALJ’s decision and federal court review. Without a realistic sense of how long changes take to get reflected in agency output, a federal judge might become further frustrated.

Finally, the Federal Judicial Center and ACUS should jointly supervise the publication of a manual describing the disability claims adjudication process in some detail and summarizing important but basic legal doctrine.⁷³⁰ This manual could orient new law clerks to issues that might otherwise prove initially confusing, such as the use of boilerplate in ALJ decisions and the applicable standard of review. If widely read, this manual would allow attorneys to file streamlined briefs that omitted repetitive background that only an inexperienced clerk might

⁷³⁰ A model from a very different context is the *Manual for Complex Litigation*, an immensely helpful text that the Federal Judicial Center supervises.

need. Although this manual would need some updating, its focus should remain on how the disability claims adjudication process works and on fundamentals of social security law. As such, it should not need constant revision or fine-tuning for particular circuits.

5. Addressing Circuit Variation

Recommendation 5. Congress should not replace the existing system of judicial review with a specialized court for social security appeals. The Appeals Council should issue opinions in a set of appeals each year that will benefit from Chevron deference and thereby reduce circuit-level variation.

Circuit boundaries account for a great deal of the inconsistency among districts.

Differences among the circuits create three chief problems. First is the challenge that inconsistency poses to the agency's administration of a single national program. The agency has responded to this challenge with its acquiescence process, whereby it instructs decision-makers to honor circuit decisions that conflict with agency policy. Otherwise, the agency attempts to interpret circuit decisions in a manner consistent with existing agency policy, and thus the decisions are not treated as offering binding legal instruction. Nine acquiescence rulings are presently in effect in the Ninth Circuit. No other circuit has more than six.⁷³¹ By ignoring all other circuit decisions, the agency risks the occasional ire of a federal judge who wonders why ALJ decisions before him repeatedly buck settled case law. But, given the modest number of circuit decisions in which the agency acquiesces, and in light of the small number of claimant wins in the courts relative to the overall number of claims the agency adjudicates, circuit precedent does not seem to be an undue threat to the agency's capacity to administer a national program consistently.

The neglect of circuit precedent that does not trigger the agency's acquiescence process, however, raises a second concern. Inconsistency among the circuits probably blunts the power

⁷³¹ See ssa.gov/OP_Home/rulings/ar-toc.html.

of case law to influence social security law and policy. Agency noncompliance with case law grows increasingly untenable if the courts of appeals are united.⁷³² From the agency's perspective, circuit disuniformity may have some value. If judges have questionable competence to devise articulation or record development requirements, the agency can better ignore judicial demands if they are confined to a single circuit. From the judiciary's perspective, however, inconsistency may weaken efforts to encourage the agency to make what judges believe to be needed changes.

Finally, inconsistency is a problem if similarly situated claimants are treated differently based solely on accident of geography. Assuming that not all variation in remand rates is due merely to variation in the quality of claims, a claimant before a federal judge in Santa Fe has a much better shot at winning a remand than her exact equivalent in El Paso, in significant measure because of circuit boundaries. This problem strikes us as the most serious one, and the one that most deserves a response.⁷³³ But it is also a hard one to address. Individual outlier judges may have incorrect views of applicable law, and appeals to outlier districts might result from problems within hearing offices. Differences among the circuits, in contrast, do not reflect anything dysfunctional but are a normal manifestation of a regional system for deciding appeals.

We do not favor two possible responses to the problem of circuit variation. The first involves the wholesale replacement of Article III review, presumably with some sort of Article I court.⁷³⁴ The arguments for and against specialized courts are many and cannot be summarized

⁷³² Cf. DERTHICK, AGENCY UNDER STRESS, *supra* note 35 at 143 (observing that the agency's uniformity-based justification for disregarding case law "came to be much weakened" by "the high degree of agreement among circuit courts").

⁷³³ *But see* MASHAW ET AL., *supra* note 28 at 111-112 (questioning whether this sort of inconsistency is normatively problematic).

⁷³⁴ *E.g.*, Verkuil & Lubbers, *supra* note 23 at 778; MASHAW ET AL., *supra* note 28 at 147-150.

concisely here,⁷³⁵ but one argument against is particularly apt. The Appeals Council, a specialized appellate body, already reviews ALJ decisions. At minimum, an additional layer of specialized appellate review would have to have a much smaller caseload to add anything distinctive. Otherwise, the specialized court might exhibit the same limitations as any other high-volume adjudicative body with a specialized docket.⁷³⁶ Claimants presently take nearly 20,000 appeals to the federal courts. Congress would have to create an enormous specialized court to ensure that appeals from the agency receive anywhere near the time and attention the district courts presently give them.⁷³⁷

A less dramatic reform would retain district court review but give the Federal Circuit exclusive appellate jurisdiction. By removing the other courts of appeals from the system, this change would ensure that only a single set of precedential decisions would govern social security litigation. We doubt, however, that the agency would find this reform an improvement. First, we are skeptical that Federal Circuit review would in fact engineer uniformity in social security case law. The Federal Circuit issues very few precedential decisions in veterans' benefits cases.⁷³⁸ This phenomenon results partly from the size of this part of its docket⁷³⁹ and partly from the limited scope of review that the Federal Circuit has for these cases.⁷⁴⁰ But it also might also reflect the Federal Circuit's felt identity as the national patent court, an identity that might

⁷³⁵ For a very useful discussion in the context of veterans' claims, see, e.g., Michael Wishnie, *Emerging Issues in Veterans Law*, at 12-23 (unpublished draft, on file with authors).

⁷³⁶ E.g., James D. Ridgeway et al., "Not Reasonably Debatable": *The Problems With Single-Judge Decisions by the Court of Appeals for Veterans Claims*, 27 STAN. L. & POL'Y REV. ____ (2015).

⁷³⁷ In FY 2014, the Court of Appeals for Veterans Claims received more than 260 appeals per judge. See Annual Report, U.S. Court of Appeals for Veterans Claims, Oct. 1, 2013, to September 30, 2014 (Fiscal Year 2014). All things held equal, a social security court with this per capita caseload, which is already quite high, would require seventy-five judges.

⁷³⁸ Paul R. Gugliuzza, *Veterans Benefits in 2010: A New Dialogue Between the Supreme Court and the Federal Circuit*, 60 AM. U. L. REV. 1201, 1252 (2011).

⁷³⁹ See, e.g., www.cafc.uscourts.gov/sites/default/files/the-court/statistics/appeals_filed_in_major_origins_10-year_06-15.pdf.

⁷⁴⁰ 38 U.S.C. § 7292(a).

relegate its non-patent docket to a second-class status.⁷⁴¹ Were the Federal Circuit to display a lack of interest in its social security docket, the paucity of appellate guidance would leave district courts to drift in their own directions. While the agency would have only a single circuit's set of decisions to deal with, inconsistency might actually worsen, as districts within the same circuit would have less legal glue holding them together.

Even if the Federal Circuit took an active interest in its social security docket, the agency and social security claimants might be worse off. While active Federal Circuit involvement might produce more uniformity among the districts, it might also produce instability in governing law.⁷⁴² The Federal Circuit constantly remakes patent law, probably in part due to its proprietary interest over the field.⁷⁴³ Were the Federal Circuit to develop the same proprietary interest in social security law, the agency would constantly have to negotiate with the court for control over key doctrines. Finally, the addition of 650 or so appeals annually to the Federal Circuit's workload would expand it by about 25%, a significant increase that would come at the expense of time spent on other cases.

We support leaving the existing system in place, as the benefits of uniformity and consistency that either alternative would promise are either too speculative or too modest to compensate for the costs of specialized review. A third alternative might lead to more uniformity and consistency without replacing the present structure. The Board of Immigration Appeals ("BIA") selects a very small number of cases out of its voluminous docket for decision by full-dress opinion, issued after consideration by a panel of adjudicators.⁷⁴⁴ Generally

⁷⁴¹ On the identity of the Federal Circuit as a national patent court, see Paul R. Gugliuzza, *The Federal Circuit as a Federal Court*, 54 WM. & MARY L. REV. 1791, 1857 (2013).

⁷⁴² We are grateful to Paul Gugliuzza for suggesting this idea.

⁷⁴³ Paul R. Gugliuzza, *Saving the Federal Circuit*, 13 CHI.-KENT J. INT. PROP. 350, 352-355 (2014).

⁷⁴⁴ In 2014, the BIA decided 28 cases by full opinion, out of 30,000 total cases completed. See www.justice.gov/eoir/precedent-decisions-volume-26; www.justice.gov/sites/default/files/eoir/pages/attachments/2015/03/16/fy14syb.pdf.

speaking, these BIA decisions receive *Chevron* deference from the federal courts,⁷⁴⁵ a doctrine designed in part to produce greater uniformity in the judicial interpretation of federal statutes.⁷⁴⁶ The Appeals Council could do the same. It could instruct analysts to look out for cases that involve uncertain or inconsistent interpretations of the Act or regulations. A qualifying appeal could be taken out of the ordinary appeals process and reviewed by a panel of AAJs, for decision by opinion.

Were the Appeals Council to act on this recommendation, one urged previously by ACUS,⁷⁴⁷ it should craft a process that would give claimant representatives and concerned organizations ample opportunity to participate in any case that might produce a precedential decision. This process should include notice of the case and its issues and an invitation to interested parties to submit amicus briefs or otherwise participate.

B. Suggestions for the Agency

The scope of our Report does not permit us to comment on everything about disability claims adjudication that we believe could improve.⁷⁴⁸ Nonetheless, we believe that the agency should consider a variety of initiatives in addition to what it currently pursues. Consistent with our charge, we limit our suggestions for improvement to aspects of agency processes that intersect with the federal courts.

Suggestion 1. The agency should investigate further the relationship between hearing office performance and work environment, on one hand, and remand rates in district courts, on the other.

⁷⁴⁵ *INS v. Aguirre-Aguirre*, 526 U.S. 415 (1999).

⁷⁴⁶ *E.g.*, Cass R. Sunstein & Adrian Vermeule, *Interpretation and Institutions*, 101 MICH. L. REV. 885, 926-927 (2003).

⁷⁴⁷ ACUS Recommendation 87-7, A New Role for the Social Security Appeals Council, 1 C.F.R. § 305.87-7(1) (1993).

⁷⁴⁸ For example, we are convinced that the eBB needs significant improvements if it will have the ameliorative effect that the agency hopes. We also believe that the How MI Doing? tool needs adjustments, especially to emphasize decisional quality more expressly.

We argue in Part IV that district-wide cultures explain a good deal of the variation in remand rates from one federal district to the next. We conclude with the hypothesis that appeals to various districts may involve ALJ decisions of uneven quality, and that these differences may explain in part why districts have different cultures. The agency should investigate this hypothesis further. As argued, we doubt that district courts will view their remand rates as aberrational or as evidence of too stringent or too lenient review unless they are convinced that the inputs for decision-making – that is, ALJ decisions – do not vary geographically.

Suggestion 2. The agency should add bottom-up, localized experiments to their quality assurance initiatives. This experimentation could include a pilot project in several hearing offices that uses district court decisions for guidance and critique.

The agency’s attempts to improve decisional quality have included a number of ventures pursued by personnel in its headquarters, including mapping out 2,000 policy compliant pathways to decisions, analysis of large amounts of data to identify recurring problems, and focused reviews that may prompt “senior managers and executives” to take corrective actions.⁷⁴⁹ These initiatives are consistent with what scholars have identified as efforts in public administration to “strongly reduce[] the scope of administrative discretion.”⁷⁵⁰ The agency should consider a different, but complementary, approach to quality assurance that scholars in another context have called “experimentalism.”⁷⁵¹ These are bottom-up, localized experiments that do not involve “a top-down form of supervision,” as Daniel Ho puts it, but rather “a method for mutual, all-around learning”⁷⁵²

⁷⁴⁹ Ray & Lubbers, *supra* note 225 at 1593; *id.* at 1598.

⁷⁵⁰ Mark Bovens & Stavros Zouridis, *From Street-Level to System-Level Bureaucracies: How Information and Communication Technology is Transforming Administrative Discretion and Constitutional Control*, 62 PUB. ADMIN. REV. 174, 177 (2002)

⁷⁵¹ See generally Michael C. Dorf & Charles F. Sabel, *A Constitution of Democratic Experimentalism*, 98 Colum. L. Rev. 267 (1998); see also Daniel E. Ho, *Does Peer Review Work? An Experiment of Experimentalism*, 69 STAN. L. REV. __ (forthcoming) (on file with authors) (discussing and testing experimentalism in the context of one type of high volume agency adjudication).

⁷⁵² Ho, *supra* note 751 at 31.

The agency could challenge hearing offices to devise their own quality assurance initiatives and thereby learn from a variety of experiments. A hearing office, for example, might use district court decisions as a basis for possible improvements. The office could prepare a memorandum regularly, perhaps every six months, that summarizes and analyzes all relevant district and circuit court opinions. This memorandum should identify trends in district court decision-making and compare the latest appellate case law to existing agency policy. ALJs should meet to discuss this memorandum, with three objectives in mind: (1) to identify problems that district courts have identified with agency decision-making; (2) to identify federal judicial preferences that can be accommodated without producing a conflict with agency policy; and (3) to identify judicial preferences that, if honored, would occasion a rupture with agency policy.⁷⁵³

ALJs should not go through this exercise in order to craft their own idiosyncratic approaches to disability determinations, or to eschew agency policy in favor of fidelity to case law. If the hearing office discovers a different but policy-compliant preference for the analysis of a particular issue – drug and alcohol materiality, for instance⁷⁵⁴ – ALJs could consider addressing the courts’ concerns going forward. But the hearing office should immediately communicate any adjustment to decision making based on case law to headquarters, to enable agency officials to confirm that the hearing office’s decisions remain policy compliant. Also, if the hearing office identifies a trend in a district court that conflicts with the agency’s approach to an issue, for example, the office should report this clash so that the agency can consider an appropriate response.

⁷⁵³ An ALJ or someone working within the hearing office should prepare this memorandum, for several reasons. While an OGC lawyer may be best situated to do so, a couple of OGC lawyers mentioned that they have encountered some resistance from ALJs when offering feedback. Moreover, such experimentalist methods of performance management are premised on peer review and evaluation. OGC Lawyer 2 at 4; OGC Lawyer 11 at 6.

⁷⁵⁴ We use this example because a federal judge told us that agency indifference to case law on this issue causes unnecessary remands. Federal Judge 19 at 6.

The agency can use this exercise or one like it as an opportunity for bottom-up reform. These memoranda and the dialogue with headquarters that they would catalyze could allow ALJs to participate in a structured conversation with the agency about best approaches to policy compliant decision-making. A common lament from ALJs is that agency headquarters does not always listen to them. This exercise could address this concern.

This proposal is just one of many such experiments or pilot projects the agency could allow in an attempt to leverage hearing office wisdom more aggressively. Our specific suggestion has a particular motivation. It comes from our interviews of ALJs working within the low remand district, who believe that this exercise, or a version of it, has improved their decision-making.

Also, our suggestion would offer the agency a way to test the educative value of case law. The agency discourages ALJs and decision writers from citing case law, and it does not typically grant RVRs based on inconsistency with case law. Only after the agency issues an acquiescence ruling will it honor a circuit opinion as binding, and district court opinions are never proper authorities. This policy makes a certain amount of sense. Were ALJs and decision writers to try to follow case law, the agency would necessarily sacrifice its legitimate pursuit of consistency in the implementation of a single national program. Case law differs from circuit to circuit, and district court decisions would add innumerable variations into the mix. Also, attempted fidelity to case law would multiply the sources that ALJs and decision writers would try to honor, amplifying the likelihood that idiosyncratic applications of the Act and regulations would produce inconsistent results.

We wonder, however, whether a bright-line rule against attention to case law sends the best signal to agency personnel about the legitimacy and educational value of federal judicial

feedback. Attention to case law and an attempt to honor federal court preferences may indeed create inconsistency and administrative burden.⁷⁵⁵ But the notion that ALJ regard for case law can seriously threaten the consistent implementation of a single national policy strikes us, as well as a couple of OGC lawyers we interviewed, as overstated,⁷⁵⁶ especially if controlled in the manner we suggest.

Decision writers should be involved in such initiatives as well. As we argue below, the work they do is centrally important to the quality of ALJ decisions. By reviewing decisions together, ALJs and decision writers can identify together what worked and what did not. An ALJ can point out recurring issues with a decision writer's drafts; a decision writer can identify recurring problems with the ALJ's instructions.

The agency stressed to us the efforts it has undertaken to improve consistency and policy compliant decision-making; it worried that our suggestion might counteract some of its efforts and encourage renegade approaches to claims adjudication. We appreciate these concerns and for this reason suggest that the agency experiment with this idea as a pilot project. It could then test whether district court opinions provide useful guidance and whether they interfere with policy compliant decision-making. Bottom-up, experimentalist approaches to quality assurance have demonstrated significant potential elsewhere.⁷⁵⁷ If not our suggestion, we encourage the agency to pursue one like it.

Suggestion 3. The Social Security Administration and the Administrative Office of the U.S. Courts should provide the federal judiciary with a database listing district and magistrate judge decision rates.

⁷⁵⁵ For example, the Fourth Circuit requires that an ALJ give "substantial weight" to the disability determination of the Veterans' Administration. *Bird v. Commissioner of Social Security*, 669 F.3d 337, 343 (4th Cir. 2012). The applicable Social Security Ruling, in contrast, insists only that VA findings "cannot be ignored and must be considered." Soc. Sec. Ruling 06-3p.

⁷⁵⁶ OGC Lawyer 23 at 3 (agreeing with this thought); OGC Lawyer 31 at 5 (insisting that ALJs could both attend to judicial preferences and honor agency policy "90%" of the time).

⁷⁵⁷ See generally Ho, *supra* note 751 (describing the encouraging results of an experiment with peer review conducted by an agency that adjudicates a high volume of issues and has struggled with consistency).

A district-wide aberrational remand rate might indicate underlying problems with ALJ decisions flowing into the court. When a single judge within a district has an aberrational remand rate, in contrast, the judge's idiosyncratic views of social security law and policy might be responsible. The luck of the draw – the judge to whom a case happens to be assigned – produces different outcomes for otherwise similarly situated claimants. Consistency, a value that a system of adjudication ought to serve, is frustrated. Moreover, consistency is a decent proxy for accuracy. If they are deciding substantial numbers of cases, rather than just a potentially non-representative handful, outlier judges are perhaps wrong on the law.⁷⁵⁸

As we reported in Part IV, judges rarely diverge all that significantly from others within their district. Still, individual judge idiosyncrasy, to the extent that it is a problem, merits a response. The availability to judges of a database that includes remand rates, broken down by judge, might help.⁷⁵⁹ Such data are available for immigration judges, for social security ALJs, and for federal judges with respect to their sentencing decisions. In the case of immigration judges, these data produced significant scrutiny and public outcry, and ultimately they prompted the Attorney General to respond. These immigration judge data have also given reviewing courts a way to calibrate their review of particular decisions.⁷⁶⁰ Some version of this benefit might result from the availability of district and magistrate judge data. The Fourth Circuit, for example, might review a district judge differently if it knew that the judge's decision pattern deviated from a district's mean.

⁷⁵⁸ Legomsky, *supra* note 443 at 425 (explaining why inconsistent adjudicators are likely incorrect).

⁷⁵⁹ A database hosted by Syracuse University compiles judge-specific sentencing data, so precedent for this sort of information gathering and publication exists. See <http://trac.syr.edu/judges/aboutData.html>.

⁷⁶⁰ For a different but related example, see *Hu v. Holder*, 579 F.3d 155, 158-59 (2d Cir. 2009). There, the Second Circuit refused to defer to an immigration judge's assessment of demeanor evidence because the IJ made findings four years after the asylum applicant testified. The Second Circuit supported this determination with data indicating how many asylum applications the IJ had ruled upon between the applicant's hearing and the issuance of his decision. *Id.* at 159.

These data should also document district-level variations. Such differences are more important than disparities in outcomes produced by outlier judges. District and magistrate judges and agency officials alike should know if a district deviates from its circuit's mean, and circuit judges should know if the mean of districts within their court's boundaries deviates from the national one. This information is a first step toward understanding what is going on within a particular district. Perhaps the problem is the jagged quality of ALJ decisions. Perhaps it is inadequate Appeals Council review. But it also might result from decisional heuristics that are misleading federal judges within a district.

Suggestion 4. The Social Security Administration should attempt to quantify the "false positive phenomenon," or the number of court remands that, once adjudicated again, do not result in the payment of benefits.

Agency officials complained to us of the cost of remands in terms of demands on agency time. But these costs mean nothing on their own. If court remands generate benefits for claimants who are in fact entitled to them, then the advantages of the more accurate implementation of disability policy may exceed the costs of judicial review. A true positive in this context is an agency denial reversed by a district court that properly results in the payment of benefits. A false positive is a claim correctly denied by the agency the first time around that a district court nonetheless remands. The value of district court review by a utilitarian metric depends, at least in part, on whether true positives exceed false positives. If so, the additional costs that remands create are probably justified. If not, a case for more deferential judicial review is stronger.

Federal judges have little idea what happens to claims on remand. The agency should gather information about ultimate outcomes and make it available to federal judges. If a federal judge remands an unusually small number of claims and all end up getting paid, she might take

another look at applicable law, to make sure that her understanding of it is not erroneous. If a colleague who remands an unusually large number of claims learns that only 25% get paid, he might do the same.

Such data would be too blunt, at least in one respect. False positives, as we have defined them, would include instances where the judge does not have any opinion as to the merits of the claim but instead remands due to some legal error in the ALJ's order. One the ALJ performs the required analysis, he may well still deny the claim. But the federal judge's decision may well have been the right one on the law.⁷⁶¹ Still, by comparing his numbers to a colleague's in the same district, a judge can get some rough sense of how many remands are really false positives. If 30% of an outlier judge's remands get paid benefits, while 70% of a colleague's remands get paid, the outlier judge might be encouraged to revisit his understanding of social security law.

Suggestion 5. To the extent possible, the Social Security Administration should require that hearing offices assign court remands to the same decision writers who worked on the cases the first time.

Decision writers are the least visible participants in the disability claims process.⁷⁶² Within the agency, of course, decision writers have well established and clear roles. But in various communications with us, agency officials repeatedly downplayed their centrality. The agency declined our request for data on decision writers' remand rates, for example, on grounds that ALJ responsibility for decisions that go out under their names renders information about decision writer performance irrelevant. We appreciate that the ALJ, not the decision writer, is the decision maker. But time pressures may preclude an ALJ's extensive review of a decision

⁷⁶¹ The agency could refine its classification scheme to separately identify cases that did not result in the payment of benefits but were nonetheless correctly remanded. Doing so would require a qualitative assessment of claims that get denied on remand, to identify those that the agency agrees were rightly remanded.

⁷⁶² One federal judge told us that he only learned of decision writers' existence twelve years and countless social security appeals into his career on the bench. Federal Judge 6 at 4. A permanent clerk for one district judge with several years of experience was unaware that they existed. Federal Judge 8 at 3. *See also* Transcript, Mar. 13, 2013, at 14, *at* www.ssaconnect.com/260-sanctions (quoting a district judge referring to decision writers as "ghostwriters").

writer's draft, and decisions therefore may reflect the writer's analysis as much as the ALJ's for certain issues.⁷⁶³ Decision writer competence is key.

The agency has long had a policy of returning a court remand to the same ALJ who decided it in the first instance.⁷⁶⁴ This policy frustrates some claimant representatives,⁷⁶⁵ but it strikes us as a useful way to ensure that ALJs receive important feedback. This policy should include decision writers as well.⁷⁶⁶ Decision writers can learn of remands from How MI Doing, but few, if any, of those whom we interviewed knew that they could get information about their performance from it. A requirement that decision writers be assigned to remands of decisions they drafted would ensure that they appreciate why their drafts failed appellate scrutiny the first time. A decision writer might not remember the draft that he wrote the first time around, having written so many in the interim. But he also could not dismiss the feedback as a comment on someone else's work, and thus he is more likely to learn from it.⁷⁶⁷

Suggestion 6. The Social Security Administration should study the issue of an OGC attorney's ethical obligations and, where appropriate, provide clearer guidance.

We detected two understandings of the OGC lawyer's role in our interviews. First, some OGC lawyers treat the agency as their client, with duties attendant to this relationship that include zealous advocacy.⁷⁶⁸ This conception came up in conversations with us about RVR practices. Although the threshold for when to request a voluntary remand differs around the

⁷⁶³ See Letter from Frank A. Cristaudo, Chief Administrative Law Judge, to Colleagues, Dec. 19, 2007, at 2 (on file with authors) ("We . . . ask the attorneys and paralegals to provide draft decisions that require little or no editing. We do not want the ALJ, or anyone else, to waste precious time re-drafting decisions.").

⁷⁶⁴ There are some exceptions to this policy. For ALJ case assignment policies, see generally HALLEX, *supra* note 38 § I-2-1-55.

⁷⁶⁵ Some believe that an ALJ who denied a claim the first time around may be particularly predisposed to doing so a second time. *E.g.*, Letter from Timothy Cuddigan & Barbara Silverstone to David Marcus, Aug. 14, 2015, at 13-14 (on file with the authors).

⁷⁶⁶ One experienced decision writer told us that she has argued for such a policy until she has been "blue in the face." Decision Writer 2 at 2.

⁷⁶⁷ A decision writer explained to us that, unless confronted with his or her own errors, a decision writer will dismiss a court remand as a comment on someone's competence. Decision Writer 2 at 4.

⁷⁶⁸ *E.g.*, OGC Lawyer 31 at 4.

country,⁷⁶⁹ several OGC lawyers we interviewed described an attitude of “defend the case, defend the case.”⁷⁷⁰ An Appeals Council analyst offered a second understanding of the role; OGC lawyers operate in part as a backstop that can catch errors that might slip past the Appeals Council. An OGC lawyer agreed with this view.⁷⁷¹ ODAR is “numbers-oriented” because of its huge volume, this lawyer told us, whereas OGC is “more quality-oriented.” When a case has a “really complicated record,” he continued, “I might have been the first person to take substantial time to look at it.”⁷⁷² Conceived of as a backstop, an OGC lawyer’s job is less the zealous advocate and involves a more neutral task.

The ethical obligations of government counsel are notoriously complicated.⁷⁷³ Based on our interviews of thirty-one OGC lawyers, we hypothesize that inconsistent understandings of what these obligations require prevail. The agency should study this issue more comprehensively, to determine if OGC lawyers indeed harbor different conceptions of their duties. This study could be the basis for more precise ethical guidance. In discussions with us, for example, the agency challenged the neutral conception of the OGC lawyer’s role, suggesting that, after multiple layers of review, the agency is entitled to have its position defended vigorously in court. If this is so, the agency should make its position explicit. Our interviews revealed that at least some personnel within the agency think differently at present.

⁷⁶⁹ OGC Lawyer 26 at 2; OGC Lawyer 2 at 2.

⁷⁷⁰ OGC Lawyer 13 at 1; OGC Lawyer 9 at 2 (noting that his office makes a policy of defending an appeal whenever it can); OGC Lawyer 7 at 3 (describing an emphasis in his office to reduce RVRs); OGC Lawyer 10 at 2 (describing how OGC lawyers would be congratulated if their RVR rates declined); OGC Lawyer 18 at 4 (describing how RVR-ing a case is not part of her mindset); OGC Lawyer 31 at 3 (describing a culture “not to RVR” and noting that he has litigated cases his “gut” says not to); OGC Lawyer 5, Second Interview at 5 (insisting that she would RVR only an “absolutely indefensible case”); OGC Lawyer 24 at 4 (insisting that she “will defend unless I have absolutely nothing to say”).

⁷⁷¹ Others did not. OGC Lawyer 18 at 4.

⁷⁷² OGC Lawyer 31 at 3. Another OGC lawyer recounted that, when he was hired, he was told that he as an OGC lawyer wore three hats – one for the agency itself, one for the taxpayer, and one for the claimant. OGC Lawyer 11 at 3. The claimant “hat” is inconsistent with an attitude of “defend the case, defend the case.”

⁷⁷³ *E.g.*, Note, *Government Counsel and Their Obligations*, 121 HARV. L. REV. 1409, 1412-1416 (2008); *Freeport-McMoRan Oil & Gas Co. v. F.E.R.C.*, 962 F.2d 45, 47 (D.C. Cir. 1992) (insisting that “government lawyers have obligations beyond those of private lawyers”).

* * *

We close with a final thought. Ultimately, all stakeholders have an interest in reducing district court remand rates, provided that the reduction results from significant improvements in the quality of agency adjudication. A variety of factors determine the rate at which claimants prevail in the district courts: claim quality; the contours of the multi-layer process of agency review; claimants' beliefs about their chances of succeeding in the district courts; the quality of claimant and agency lawyering; variation in applicable circuit law; and judicial preferences and choices. In the presence of such a complex set of interactions, it is difficult to say with any confidence what impact on the claimant win remand rate any particular reform would have. We nonetheless believe that only a dramatic reduction in ALJ caseloads could permit significant, across-the-board improvements in decision-making quality sufficient to cause the federal court remand rate to plummet sharply. To avoid a spike in the backlog of claims, the size of the ALJ corps would have to increase. Ultimately, this may be the most important reform of all.