Procedural Roulette

Amy E. Semet[1]

In *Refugee Roulette*, Jaya Ramji-Nogales, Andrew Schoenholtz, and Philip Schrag showed that the outcome of an asylum case often depends less on the facts of the claim than on the identity of the judge who hears it, with grant rates ranging from near zero to almost 100% among judges sitting in the same building and reviewing applications from the same country. That finding launched two decades of research into adjudicator disparity, nearly all of which has treated the disparity as a disagreement about the merits, traceable to a judge's ideology, professional background, or reading of immigration law.

This Article returns to *Refugee Roulette* and asks whether part of the roulette is procedural rather than substantive. Immigration judges do not only decide cases; they manage them. In managing them, they make a stream of procedural choices that shape a case long before anyone reaches the merits. Drawing on more than sixteen million proceedings decided between 2001 and 2025 in the database maintained by the Executive Office for Immigration Review, I measure six dimensions of that procedural discretion at the level of the individual adjudicator: access to counsel, access to adjudication, the timing of that access, procedural termination, record development, and adjudicative attention.

I find that immigration judges differ substantially and persistently across all six dimensions, that these differences survive adjustment for case mix, court, and time and are not explained by the strength of the cases a judge happens to hear, and that a judge's procedural practice is only weakly correlated with her overall tendency to grant relief. The central finding follows when these results are joined to outcomes: much of the between-judge variation in case results that scholarship has attributed to disagreement about the merits is instead associated with how judges manage process. Some of the refugee roulette, in other words, is a procedural roulette. After establishing this, I take up the mechanisms through which process moves outcomes and set out reforms directed at the institutions that create procedural discretion and could constrain it.

Introduction

Do immigration judges differ not only in how they decide cases, but in how much process they provide before deciding them? The first half of that question has a settled answer. For two decades,

[1] Associate Professor of Law, University at Buffalo School of Law. I thank workshop participants and the librarians who assembled the underlying records. The data construction follows the approach I developed in earlier work on the same EOIR database, extended here from outcomes to process. I will upload a detailed Appendix on how I constructed the database from EOIR records, which is a complicated task. All errors are my own.

scholars have documented that the outcome of an immigration case depends heavily on the identity of the judge assigned to hear it.[2] Asylum grant rates that range from near zero to near total across judges sitting in the same building, hearing demographically similar applicants, have made "refugee roulette" a familiar shorthand for a system in which the law on the books promises uniformity and the law in practice delivers something closer to a lottery.[3] The second half of the question, whether judges differ in the process they provide, has gone largely unasked.

This Article addresses it directly using a comprehensive database drawn from government records. Immigration judges do far more than grant or deny relief. They manage dockets of extraordinary size under conditions of chronic backlog, and in doing so they make a continuous stream of procedural choices that determine how a case unfolds long before anyone reaches the merits.[4] A judge decides whether to grant a continuance so that a respondent can find a lawyer, gather country-conditions evidence, or await the adjudication of a collateral application. A judge decides how many master calendar and individual hearings a case will receive, and how far apart to set them. A judge decides how long a case will remain open. A judge decides, when a respondent fails to appear, whether to order removal in absentia or to give the respondent another chance to explain.[5] None of these choices is the merits. All of them shape the merits.

[2] The foundational study is Jaya Ramji-Nogales, Andrew I. Schoenholtz & Philip G. Schrag, *Refugee Roulette: Disparities in Asylum Adjudication*, 60 STAN. L. REV. 295 (2007). For a more recent treatment using the same data

[3] See Ramji-Nogales et al., *supra* note 2, at 296 (reporting that, in one regional asylum office, applicants' chances of success varied by more than fifty percentage points depending on the officer assigned).

[4] On the scale of the backlog, see HOLLY STRAUT-EPPSTEINER, CONG. RSCH. SERV., IN12492, FY2024 EOIR IMMIGRATION COURT DATA: CASELOADS AND THE PENDING CASES BACKLOG 1 (2025) (reporting more than 3.6 million pending cases).

[5] The governing regulation provides that an immigration judge "may" order removal in absentia upon a showing that written notice was provided and that the noncitizen is removable, 8 U.S.C. § 1229a(b)(5)(A), but the decision to proceed in absentia, to continue, or to reset depends heavily on the individual judge. See Ingrid V. Eagly & Steven Shafer, *A National Study of Access to Counsel in Immigration Court*, 164 U. PA. L. REV. 1 (2015).

I call the sum of these choices a judge's procedural style, and I call the authority to make them procedural discretion. The term is borrowed, with modification, from recent administrative law scholarship. Bijal Shah has argued that agencies engaged in informal adjudication exercise a form of discretion that operates not on outcomes but on process, on notice, on the opportunity to be heard, on whether a decision is made on a record at all, and that this "procedural discretion" escapes the attention of courts and scholars precisely because it does not look like a decision on the merits.[6] Shah develops the concept across several agencies, including the immigration agencies, and argues that procedural discretion tends to reduce process, that it varies widely both across and within categories of adjudication, and that it falls hardest on the people least able to absorb it.[7] Shah's account is conceptual and institutional. It identifies the phenomenon, locates its legal sources, and explains why it matters. What it does not do, what no one has done, is measure it.

That is the task of this Article. Procedural discretion is not merely a feature of agencies in the abstract; in immigration court it is exercised by individual adjudicators, case by case, and it leaves a record. Every continuance, every hearing, every reset, every in absentia order is logged in the administrative database maintained by the Executive Office for Immigration Review

source relied on here, see Catherine Y. Kim & Amy Semet, *An Empirical Study of Political Control over Immigration Adjudication*, 108 GEO. L.J. 579 (2020).

[6] Bijal Shah, *Administrative Procedural Discretion*, 111 CORNELL L. REV. (forthcoming 2026) (manuscript at 101) (defining procedural discretion as "administrative discretion that impacts the contours and defensibility of administrative process in informal adjudication").

[7] *Id.* (manuscript at 102) (arguing that procedural discretion "tends to limit the process underlying informal administrative adjudication" and that "under-resourced communities . . . are more likely to be harmed by procedural discretion").

("EOIR"), the Department of Justice ("DOJ") subcomponent that houses the immigration courts. That database contains more than sixteen million proceedings drawn from millions of cases, each tied to a specific judge, a specific court, and a specific sequence of procedural events.[8]

In this Article, I analyze EOIR proceeding-level records from 2001 through 2025, and I construct judge-level measures of six dimensions of procedural discretion: access to counsel, measured by counsel-seeking continuances and by the rate at which initially unrepresented respondents become represented; access to adjudication, measured by the rate at which a judge's cases reach an individual merits hearing; the timing of that access, measured by the master calendar hearings and the elapsed time that precede the first merits hearing; procedural termination, measured by the share of a docket resolved by in absentia order, abandonment, withdrawal, or other procedural default rather than on the merits; record development, measured by the continuances a judge grants so that a respondent can gather evidence and prepare a claim; and adjudicative attention, measured by the number and complexity of hearings a case receives once it reaches the merits. I treat these not as case characteristics but as judge characteristics, attributes of the adjudicator that persist across the cases assigned to her. I then ask three questions. First, do immigration judges differ systematically in procedural style, or does procedural variation simply track the cases that happen to land on a given docket? Second, if judges do differ, are those differences stable, does a judge who provides more process in one year provide more process in

[8] The data are released periodically by EOIR in response to standing FOIA requests and are mirrored by the Transactional Records Access Clearinghouse at Syracuse University. For a description of an earlier version of the same database, see Kim & Semet, *supra* note 2, at 600–04.

the next? Third, does procedural style predict outcomes, such that respondents before high-process judges fare differently than otherwise similar respondents before low-process judges?

The structure of the inquiry marks a departure from the existing literature. That literature treats the judge as a bundle of preferences over outcomes, a disposition to grant or deny, and measures the judge by the rate at which she grants or denies.[9] This Article treats the judge as something more: a manager of process whose procedural choices are themselves a dimension of adjudicatory behavior, distinct from and prior to the disposition. A judge can be generous on the merits and stingy with process, or the reverse. The two need not move together, and whether they do is an empirical question rather than an assumption.

The stakes of that question reach beyond immigration law. Administrative law has long debated how much process informal adjudication requires, and the doctrinal answer, since *Mathews v. Eldridge*, has been "it depends", on the private interest, the risk of error, and the government's burden.[10] But the *Mathews* calculus assumes that the relevant procedural choices are made at the level of the agency or the program, through rules that apply across cases.[11] It has little to say about a system in which the amount of process a person receives depends not on a rule but on the adjudicator she draws. If procedural discretion is real, persistent, and consequential at the level of the individual judge, then the due process question is not only how much process the system provides on average but how unevenly it provides it, and that is a question *Mathews* was not designed to answer.[12]

[9] See, e.g., Ramji-Nogales et al., *supra* note 2; Kim & Semet, *supra* note 2; BANKS MILLER, LINDA CAMP KEITH & JENNIFER S. HOLMES, IMMIGRATION JUDGES AND U.S. ASYLUM POLICY (2015).

[10] Mathews v. Eldridge, 424 U.S. 319, 335 (1976).

[11] See *id.* at 344–47 (evaluating the adequacy of procedures governing the termination of disability benefits as a categorical matter).

[12] For the argument that *Mathews* balancing has crowded out attention to the distribution of process, see Shah, *supra* note 6 (manuscript at 102–04); cf. Jerry L. Mashaw, *The Supreme Court's Due Process Calculus for*

I preview the principal findings here and develop them at length below. Immigration judges differ markedly across all six dimensions. After adjustment for case mix, court, and time, a judge at the 90% of the access-to-counsel distribution grants counsel-seeking continuances at more than three times the rate of a judge at the 10%, and her initially unrepresented respondents go on to obtain counsel far more often. A judge at the 90% of the access-to-adjudication distribution brings three-fifths of her cases to a merits hearing where a judge at the 10% brings barely a quarter. Judges also differ by hundreds of days in how quickly they move a case to its first merits hearing and by tens of percentage points in how much of the docket they resolve by procedural termination. These differences are not noise, and they are not an artifact of the cases a judge draws. They persist when I compare judges in the same court in the same period.

The finding that matters most follows when these results are joined to outcomes. A onestandard-deviation increase in the assigned judge's procedural index is associated with a relief rate

---

roughly 6% points higher, and a procedural-termination rate roughly 10% lower, for otherwise similar respondents. Not surprisingly, the associations are largest among the unrepresented and the detained, the respondents least able to supply for themselves the process a parsimonious judge withholds. Decomposing the between-judge variation in case outcomes into a part associated with procedural management and a part associated with merits disposition, I tentatively (at this time) attribute between a quarter and a third of that variation to procedural management. A meaningful portion of what two decades of scholarship since *Refugee Roulette* have called variation in how

---

Administrative Adjudication in Mathews v. Eldridge: Three Factors in Search of a Theory of Value, 44 U. CHI. L. REV. 28 (1976).

immigration judges decide is, on the most defensible reading of the data, variation in how they manage process. Some of the refugee roulette is a procedural roulette.

Three contributions follow. The first is empirical. Administrative-law scholars, most recently Bijal Shah, have argued that adjudicators in informal systems exercise a form of discretion that operates on process rather than on outcome, and that this procedural discretion is wide, uneven, and consequential. That argument has been developed conceptually. This Article does not propose a new theory of administrative procedure; it supplies the evidence that an existing one has lacked, showing that procedural discretion can be measured, that it varies systematically across adjudicators, that the variation is stable, and that it is associated with the outcomes the system exists to produce. The second contribution is methodological. The measures developed here are built from administrative event records rather than from coded merits decisions, and the approach transfers to other mass-adjudication systems, from Social Security disability hearings to veterans' claims, wherever adjudicators manage process at volume and the management leaves a record. The third is institutional. The findings bear on live debates about case-completion quotas, the continuance standard, administrative closure, and whether the immigration courts should be removed from the Department of Justice.

Because the central claim is that procedural management matters independently of the cases judges hear, the analytical burden falls on separating the two, and I treat that burden as the core of the Article rather than as a methodological appendix. The obvious objection to every finding reported here is that process is responding to case strength rather than shaping outcomes. Stronger claims may require more hearings and more time; represented respondents may both attract more process and fare better; complex cases may stay open longer whatever judge hears them. If that is all that is happening, there is no procedural judge, only a procedural reflection of the docket. Part IV is built around defeating that objection, through judge fixed effects, court-byperiod

comparisons, leave-one-out judge measures, restriction to rotation-like assignment, and analyses that hold representation and claim type constant. The findings survive each, which is why I state them as findings about judges rather than about cases.

Procedural variation has stayed out of view for a reason worth naming, because the reason explains why measuring it is harder than measuring outcomes and why the effort is overdue. A grant or a denial is a salient, singular event: it has a date, a valence, and a person who won or lost, and it is therefore easy to count and easy to compare. Procedural management is diffuse. It is spread across a sequence of small choices, a continuance here, a reset there, an extra hearing, an earlier default, none of which looks momentous on its own, and the cumulative pattern emerges only when the choices are gathered across a judge's whole docket. One needs detailed records, as the EOIR has supplied, to even analyze it. The disposition announces itself; the procedural style must be reconstructed from those detailed records. That asymmetry, and not any judgment that process matters less than outcome, is why the literature has measured the one and not the other.

This Article reconstructs the procedural style, and in doing so makes visible a form of judicial variation that has been hiding in plain sight, recorded in the data the whole time but never assembled into a picture of how judges differ.

This Article proceeds in five Parts. Part I joins two literatures: the empirical literature on adjudicator disparity that *Refugee Roulette* began, and the administrative-law literature on procedural discretion in informal adjudication. Part II sets out the six dimensions of procedural discretion and the judicial choices within each that the records make visible. Part III describes the EOIR database and the construction of the judge-level measures, following the approach I developed in earlier work on the same source. Part IV is the empirical core as it explains the identification strategy at length, shows that judges differ across the six dimensions and that the

differences are not an artifact of case selection, and relates procedural practice to outcomes. Part V takes up the mechanisms through which process affects outcomes.

I. Two Literatures, Unjoined

The claim that immigration judges differ is not new. The claim that they differ in their procedural practice, and that this difference is worth measuring on its own terms, is. This Part explains why by tracing two bodies of scholarship that have developed in parallel without meeting. The first is the empirical literature on outcome variation in immigration adjudication, which has established beyond serious dispute that who decides a case shapes how it comes out. The second is the administrative law literature on procedural discretion, which has recently begun to take process itself as the object of study. This literature has remained conceptual and has not descended to the level of the individual adjudicator or the measurable case. Joining the two yields the question this Article pursues.

A. What We Know About How Immigration Judges Decide

The modern study of immigration adjudication begins with disparity. In *Refugee Roulette*, RamjiNogales, Schoenholtz, and Schrag examined asylum decisions across immigration courts and asylum offices and found differences so large that they could not plausibly be explained by the characteristics of applicants.[13] In the same court, hearing applicants from the same country, one judge might grant asylum in nine of ten cases and another in one of ten.[14] The authors framed the finding as a rule-of-law problem: a system that produces such divergence on identical facts is not applying law so much as expressing the predispositions of the people who staff it.[15] The phrase caught on because the image was apt, and because the data were hard to argue with.

---

[13] Ramji-Nogales et al., *supra* note 2, at 295–96. The authors examined more than 140,000 asylum decisions.

[14] *Id.* at 333–35 (documenting grant-rate ranges within single immigration courts).

[15] *Id.* at 378–79 (describing the disparities as a threat to the perceived legitimacy of the asylum system).

The work that followed refined the picture without disturbing its central claim. Scholars connected outcome variation to judges' professional backgrounds, asking whether judges who had previously worked as prosecutors or as immigration enforcement attorneys ruled differently from those who had represented noncitizens or worked in nongovernmental organizations.[16] Others related variation to gender, to caseload, to the time of day or the cumulative number of grants a judge had already issued, drawing on behavioral findings about decision fatigue.[17] A distinct strand turned from the attributes of the judge to the political environment in which she worked, asking whether immigration judges, who are DOJ employees rather than Article III judges, respond to the priorities of the administration that supervises them. I, along with Catherine Kim, using the same EOIR database analyzed here, examined more than 830,000 removal proceedings and found that

---

the appointing administration of a judge and the administration in power at the time of decision both predicted removal outcomes, even after controlling for an array of case characteristics.[18] Our finding, that decisions in a purportedly adjudicative body track the political branches, reframed disparity not as idiosyncrasy but as a symptom of political control over adjudication.[19]

A second body of work, closer to this Article's concerns, looks at the process surrounding the immigration judge's decision rather than on the decision itself. Ingrid Eagly and Steven Shafer's national study of access to counsel showed that representation is at once powerfully

---

[16] See Miller, Keith & Holmes, *supra* note 9, at 60–85 (relating asylum outcomes to immigration judges' prior professional experience); see also ANDREW I. SCHOENHOLTZ, JAYA RAMJI-NOGALES & PHILIP G. SCHRAG, LIVES IN THE BALANCE: ASYLUM ADJUDICATION BY THE DEPARTMENT OF HOMELAND SECURITY (2014).

[17] For the behavioral line of inquiry, see, e.g., Daniel L. Chen & Jess Eagel, *Can Machine Learning Help Predict the Outcome of Asylum Adjudications?*, PROC. 16TH INT'L CONF. ON A.I. & L. 237 (2017); cf. Shai Danziger, Jonathan Levav & Liora Avnaim-Pesso, *Extraneous Factors in Judicial Decisions*, 108 PROC. NAT'L ACAD. SCI. 6889 (2011).

[18] Kim & Semet, *supra* note 2, at 583–84, 620–28. The authors found, among other things, that immigration judges appointed under Republican Attorneys General ordered removal at higher rates, and that removal rates shifted with changes in presidential administration.

[19] *Id.* at 629–33. For the broader theoretical claim, see Catherine Y. Kim, *The President's Immigration Courts*, 68 EMORY L.J. 1 (2018).

consequential and wildly uneven, varying with custody, court, and region. Eagly's separate work on remote adjudication showed that the format of a proceeding, who appears in person and who appears by video, shapes participation and outcomes.[20] Others have studied detention, notice, and the lived experience of the process. Emily Ryo's work on detention and bond hearings established that detained respondents are systematically less able to participate, to secure counsel, and to obtain relief, and that whether respondents experience the process as fair shapes their compliance and their view of the system's legitimacy. Fatma Marouf has catalogued the procedural irregularities that pervade removal proceedings, from inadequate competency protections to uneven evidentiary practice. Lindsay Nash has examined the notice and in absentia practices that determine whether a respondent is heard at all. Further, Jill Family and Shoba Sivaprasad Wadhia

have each situated immigration adjudication within administrative law, the first by treating the immigration courts as a study in the procedural thinness of informal adjudication, the second by tracing how low-visibility discretion, exercised at scale, governs who is removed and who is spared.[21] What unites this work is its insistence that process, representation, detention, notice, the opportunity to be heard, is a determinant of the merits decision and not a backdrop to it. What none of it has done is measure how the provision of that process varies across the individual judges who control it. That is the gap this Article addresses.

[20] On the relationship between continuances and the practical ability to obtain counsel, see *id.* at 47–49; see also Ingrid V. Eagly, *Remote Adjudication in Immigration*, 109 NW. U. L. REV. 933, 970–74 (2015) (describing how procedural format affects representation and participation); Ingrid V. Eagly, Steven Shafer & Jana Whalley, *Detaining Families: A Study of Asylum Adjudication in Family Detention*, 106 CALIF. L. REV. 785 (2018); Jill E. Family, *Due Process and the Immigration Courts*, 2021 WIS. L. REV. 837. The counsel-seeking continuance is, for this reason, the single procedural choice most tightly linked to substantive outcome, and Part II.F treats it as a measure in its own right.

A related body of work has examined the stages of the system that surround the immigration judge's decision. Studies of the Board of Immigration Appeals have documented how the streamlining reforms of the early 2000s, which permitted single-member affirmances without opinion, compressed appellate review and shifted weight back onto the trial-level decision.[22] Others have studied detention and its effect on the practical capacity of respondents to participate, showing that detained respondents secure counsel far less often and obtain relief far less often than those who are free.[23] This work matters to the present project because it identifies the populations, namely, the unrepresented and the detained,whose appeals are compressed or never taken, for whom the immigration judge's procedural choices are least likely to be corrected by anything downstream. For these respondents, the trial-level allocation of process is, in practical effect, the whole of the process they will receive, which is precisely why measuring how that allocation varies across judges is worth the effort.

[21] See Eagly & Shafer, *supra* note 5, at 32–46; Emily Ryo, *Detained: A Study of Immigration Bond Hearings*, 50 LAW & SOC'Y REV. 117 (2016); Emily Ryo, *Fostering Legal Cynicism Through Immigration Detention*, 90 S. CAL. L. REV. 999 (2017); Emily Ryo & Ian Peacock, *The Landscape of Immigration Detention in the United States* (Am. Immigr. Council 2018). On procedural irregularities and the structural features of the immigration courts, see Fatma E. Marouf, *Alternative Bases for Immigration Detention*, 17 NEV. L.J. 503 (2017); Fatma E. Marouf, *Becoming Unconventional: Constricting the Particular Social Group Ground for Asylum*, 44 N.C. J. INT'L L. 487 (2019); Lindsay Nash, *Universal Representation*, 87 FORDHAM L. REV. 503 (2018); Lindsay Nash, *Re-Examining In Absentia Removal*, 92 FORDHAM L. REV. 1 (2023); Jill E. Family, *A Broader View of the Immigration Adjudication Problem*, 23 GEO. IMMIGR. L.J. 595 (2009); SHOBA SIVAPRASAD WADHIA, BANNED: IMMIGRATION ENFORCEMENT IN THE TIME OF TRUMP (2019); SHOBA SIVAPRASAD WADHIA, BEYOND DEPORTATION: THE ROLE OF PROSECUTORIAL DISCRETION IN IMMIGRATION CASES (2015). This Article shares these scholars' premise that the distribution of process is itself a determinant of outcomes; it differs in measuring that distribution at the level of the individual adjudicator.

[22] On the Board's streamlining reforms and their effect on the quality and visibility of appellate review, see John R. B. Palmer, Stephen W. Yale-Loehr & Elizabeth Cronin, *Why Are So Many People Challenging Board of Immigration Appeals Decisions in Federal Court?*, 20 GEO. IMMIGR. L.J. 1 (2005); see also Lenni B. Benson, *Making Paper Dolls: How Restrictions on Judicial Review and the Administrative Process Increase Immigration Cases in the Federal Courts*, 51 N.Y.L. SCH. L. REV. 37 (2006). [23] See *supra* note 21.

Across these studies, two methodological commitments recur, and both are worth naming because this Article departs from them. First, the dependent variable is almost always the disposition: grant or deny, relief or removal. The literature measures the judge by the rate at which

she reaches a particular merits outcome, and it treats everything upstream of that outcome, the continuances, the hearings, the resets, as either noise or as a control variable to be partialed out.[21] Second, where the literature does attend to procedure, it does so to explain outcomes, not to study procedure for its own sake. Representation by counsel is the leading example. Eagly and Shafer's national study of access to counsel showed that represented respondents fared dramatically better than unrepresented ones, and that representation rates varied enormously across courts and custody settings.[22] That work frames representation as an input to the merits decision, a thing that improves a respondent's odds, rather than as the product of procedural choices that judges themselves influence, for instance by granting or denying the continuances that allow a respondent time to find a lawyer.[26]

The result is a literature rich in findings about how immigration judges decide and nearly silent about how they manage. We know that grant rates vary by judge. We do not know whether continuance rates vary by judge, whether the variation is as large, whether it is as persistent, or whether it is correlated with the grant-rate variation everyone has studied. We know that representation predicts outcomes. We do not know whether judges differ systematically in the procedural conditions that make representation possible. These are not gaps at the margin. They concern the bulk of what an immigration judge actually does, day to day, because most of a judge's

[21] The asylum-disparity studies typically restrict attention to completed, merits-adjudicated cases and exclude cases resolved on procedural grounds. See Ramji-Nogales et al., *supra* note 2, at 305–08 (describing case-selection criteria). This is sensible for studying merits decisions but it defines procedural management out of the analysis.

[22] Eagly & Shafer, *supra* note 5, at 2 (finding that only 37% of all immigrants, and 14% of detained immigrants, secured representation, and that represented immigrants obtained more favorable outcomes at every stage). [26] See *supra* note 20.

docket time is consumed not by merits adjudication but by the management of cases that are continued, reset, consolidated, severed, and carried over from one calendar to the next.[23]

B. Procedural Discretion in Administrative Adjudication

While the empirical literature on immigration judges was maturing, a separate conversation was unfolding in administrative law about the procedures that govern adjudication outside the formal, APA-governed model. Most federal adjudication is informal.[24] It is not conducted by administrative law judges under the formal-hearing provisions of the Administrative Procedure Act; it is conducted by a heterogeneous corps of hearing officers, agency adjudicators, and, in immigration, immigration judges who are DOJ attorneys exercising delegated authority.[25] Informal adjudication has no defined procedural floor of the kind the APA supplies for formal hearings, and the procedures that do exist come from a patchwork of statutes, regulations, agency manuals, and unwritten practice.[26]

Into this setting Shah introduces the concept of procedural discretion, with three components relevant here. First, procedural discretion is discretion over process rather than over outcome: it concerns "the contours and defensibility of administrative process," including notice, the opportunity to testify, and whether a decision is made on a record.[27] Second, procedural discretion is decentralized and variable. Because no statute fixes the procedures of informal

[23] The structural dominance of case management over merits adjudication is evident in the proceeding counts themselves: the EOIR database records many more hearing and scheduling events than completed merits decisions. See *infra* Part III.

[24] See MICHAEL ASIMOW, ADMINISTRATIVE LAW JUDGES AND HEARING OFFICERS: A JOB DESCRIPTION (2019); see also Kent Barnett, *Against Administrative Judges*, 49 U.C. DAVIS L. REV. 1643 (2016).

[25] Immigration judges are appointed by the Attorney General and are not ALJs within the meaning of 5 U.S.C. § 3105. See 8 C.F.R. § 1003.10; see also Kim, *supra* note 19, at 9–14.

[26] See Shah, *supra* note 6 (manuscript at 102) (observing that "[i]n informal adjudication there is no defined procedural floor, creating wild variations in process").

[27] *Id.* (manuscript at 101).

adjudication, agencies and the adjudicators within them can expand or contract process, producing "great variation among the processes underlying informal adjudication, not only across different types of adjudication, but also within the same category of adjudication."[28] Third, procedural discretion is distributively skewed. It tends to reduce process, and the reduction lands on "underresourced communities, and those in need of governmental support or protection," while leaving the better-resourced comparatively untouched.[29] Shah locates immigration squarely within the account, describing the discretion that immigration judges enjoy to consolidate or sever cases, to hold hearings by video, to decide claims orally or in writing without formal findings, to control whether a representative may even be present at certain hearings, and to determine how long a noncitizen has before being found to have defaulted.[30]

The conceptual payoff of Shah's framing is to make process visible as a site of discretion, and therefore as a site of potential unfairness, that the standard tools of administrative law do not police. Judicial review fixes on outcomes and on the reasons given for them; it rarely reaches the antecedent question of how much process the adjudicator chose to provide.[31] The due process doctrine that might seem to address process directly has, since *Mathews*, evaluated procedures categorically and deferentially, asking whether a class of procedures is adequate on balance rather than whether the procedures actually delivered in a particular case were the product of unreviewed

---

[28] *Id.* (manuscript at 101–02).

[29] *Id.* (manuscript at 102).

[30] *Id.* (manuscript at 143–45) (cataloguing immigration judges' procedural discretion under the Immigration Court Practice Manual and 8 C.F.R. § 1240).

[31] *Id.* (manuscript at 102) (noting that "neither procedural discretion itself nor the outcomes of decisions based in discretionary administrative procedure are necessarily subject to judicial review").

adjudicator choice.[32] Procedural discretion thus occupies a doctrinal blind spot: too granular for due process balancing, too procedural for ordinary review of the merits.

What Shah's account establishes by argument, this Article tests by measurement. Shah's procedural discretion is an institutional property, a feature of agencies and of the legal regime governing informal adjudication. To study it empirically in immigration court, I relocate it to the level of the individual adjudicator and to the level of the case. Procedural discretion, in the form I measure, is the variation in process that immigration judges actually produce across the cases assigned to them: the differences in continuance practice, hearing intensity, case duration, and default reliance that remain after the case mix and the court are held constant. If Shah is right that procedural discretion generates wide variation within the same category of adjudication, that variation should be visible in the EOIR records as systematic differences among judges. If she is right that procedural discretion is not merely random but patterned and consequential, those differences should persist over time and should bear on outcomes. The concept, in other words, generates testable predictions, and the immigration courts supply an unusually good place to test them.

C. Procedural Style as a Judicial Trait

The claim that a judge has a procedural style, distinct from her substantive disposition, draws on two literatures outside immigration law that the immigration scholarship has not engaged. The first is the social-psychological work on procedural justice, associated with Tom Tyler and others, which holds that people's sense of whether they have been treated fairly depends heavily on the process they experience, whether they were heard, whether the decisionmaker appeared neutral,

[32] Mathews, 424 U.S. at 335; see also Shah, *supra* note 6 (manuscript at 103–05).

whether the procedure conveyed respect, and not only on whether they won.[33] If process carries independent weight in how people perceive the legitimacy of an adjudication, then variation in process is a fairness problem in its own right, not merely a proxy for variation in outcomes. A respondent who is rushed through a single hearing, denied the time to find a lawyer, and removed in absentia has been treated differently, and, on the procedural-justice account, less fairly, than one afforded continuances and a full merits hearing, even if a reviewing court would have reached the same result on the merits in both cases. The procedural-justice literature thus supplies a reason to care about procedural variation that does not depend on showing that process changes outcomes.

The second is the political-science literature on judicial behavior and case management, which has long recognized that judges differ not only in how they vote but in how they run their courtrooms.[34] Studies of federal district judges have examined variation in how quickly judges move their dockets, how readily they grant extensions, and how aggressively they manage pretrial proceedings, treating these as stable features of a judge's approach to the work rather than as random noise. That literature has not, however, been carried into the immigration courts, where the management function is, if anything, more central to the judge's role than in an Article III court, because the immigration judge presides over a high-volume docket with few of the procedural

[33] See TOM R. TYLER, WHY PEOPLE OBEY THE LAW (1990); Tom R. Tyler, *Procedural Justice, Legitimacy, and the Effective Rule of Law*, 30 CRIME & JUST. 283 (2003). The procedural-justice literature establishes that the experience of process carries independent weight, which is one reason variation in process is normatively significant apart from its effect on outcomes.

[34] On trial-judge case-management variation, see, e.g., STEVEN FLANDERS, CASE MANAGEMENT AND COURT MANAGEMENT IN UNITED STATES DISTRICT COURTS (1977); Marc Galanter, *The Vanishing Trial: An Examination of Trials and Related Matters in Federal and State Courts*, 1 J. EMPIRICAL LEGAL STUD. 459 (2004). The study of how judges manage cases, as opposed to how they decide them, is well developed for Article III trial courts but has not been carried over to immigration adjudication.

rules that structure ordinary civil litigation. The immigration judge is, to a degree unusual among American adjudicators, a manager first and a decider second.

Bringing these two literatures together with the immigration scholarship yields the concept this Article measures. From the procedural-justice work comes the premise that process has value independent of outcome, so that its uneven distribution is a harm worth identifying. From the casemanagement work comes the premise that procedural practice is a stable, measurable attribute of an individual adjudicator, not a byproduct of the cases she happens to draw. From the immigration scholarship comes the setting, a high-volume adjudicative system, staffed by individual judges, that records what those judges do. Procedural style is what sits at the intersection: a judge-level trait, expressed through the management of process, that the records make visible and that the existing immigration literature, fixed on disposition, has passed over.

One implication of this framing deserves emphasis because it shapes the empirical design. If procedural style is a genuine trait rather than a reflection of case mix, it should behave like other measured traits: it should be stable within a judge over time, it should distinguish judges facing similar caseloads, and it should not collapse into the substantive disposition with which the prior literature has been preoccupied. Each of these is a testable proposition, which Part IV addresses. The conceptual claim that judges have procedural styles is, in the end, only as good as the evidence that the styles persist, that they survive adjustment for what judges hear, and that they are not merely substantive generosity wearing a procedural disguise.

D. Why Immigration Court Is the Right Place to Look

Three features make the immigration courts a setting in which procedural discretion can be observed at scale and observed cleanly enough to separate from confounds. None of the three is

fully present in most other adjudicative systems, which is part of why procedural discretion has remained an argument rather than a finding.

The first is volume and recordkeeping. EOIR adjudicates an enormous number of cases and logs them in a relational database designed for case management, which means that procedural events, the scheduling of a hearing, the granting of a continuance and its coded reason, the entry of an in absentia order, are recorded as a matter of administrative routine rather than research design.[35] The data were generated to run the courts, not to study them, which makes them an unusually direct trace of what judges did rather than what they reported doing. Because the data is already collected, the researcher does not have to conduct the painstaking duty of compiling the records themselves from incomplete data.

The second is the assignment structure. Within a given immigration court, cases are distributed among the judges sitting there, and although assignment is not random in the strict experimental sense, it is governed by docketing and scheduling practices that are not tied to the procedural questions this Article studies.[36] That structure makes it possible to compare judges who face broadly similar pools of cases, and to ask whether their procedural practice diverges even when the underlying caseload does not.

The third is the separability of process from merits. In immigration court, the procedural events are recorded independently of the disposition. A case can be continued five times and then granted; it can be continued once and then denied; it can be resolved in absentia without any merits adjudication at all. Because the database captures the procedural sequence and the outcome as

[35] See *infra* Part III (describing the structure of the EOIR database and the proceeding-level event records). For the scale of EOIR's docket, see STRAUT-EPPSTEINER, *supra* note 4, at 1.

[36] The non-randomness of assignment and the steps taken to address it are discussed *infra* Part IV. For analogous treatment in the asylum context, see Ramji-Nogales et al., *supra* note 2, at 346–47.

distinct fields, it is possible to measure a judge's procedural style without reference to how she rules on the merits, and then to relate the two. In most studies of judicial behavior, by contrast,

process and outcome are observed together or not at all, and procedure can be inferred only from the decision it produced.

These features do not make immigration court a laboratory. The assignment of cases is imperfectly controlled and EOIR data carry the errors and idiosyncrasies of an administrative system under strain. Further, the population of cases changes over a quarter century of shifting law and enforcement priority. Part IV confronts each of these problems directly rather than assuming it away. But the unique nature of EOIR data means that if procedural discretion exists as a measurable feature of adjudicator behavior, it should be detectable here. The remainder of this Article asks whether it is.

II. Six Dimensions of Procedural Discretion in Immigration Court To measure procedural discretion, one must first say what it is made of, and the answer should come from a view of what process does for a respondent rather than from whatever the data happen to record. I organize procedural discretion around six dimensions, and they are not an arbitrary list. They track the path a removal case takes from its opening to its close, and each marks a point at which a judge's discretion determines how much process the respondent receives. Before a case can be decided well, the respondent must be able to obtain a lawyer; the case must reach a hearing on the merits; it must reach that hearing without undue delay; it must not be diverted to a procedural exit that ends it without a merits decision; the record on which the decision rests must be allowed to develop; and the hearing itself must receive adequate adjudicative attention. The six dimensions,

access to counsel, access to adjudication, the timing of that access, procedural termination, record development, and adjudicative attention, are the joints of that sequence, and a judge can extend or compress process at each.

The framework is ideal, because it tells us which procedural choices belong in the measures and which do not. The choices I measure are the ones that are at once central to the path of a case, exercised by the individual judge rather than fixed by rule, variable across judges, and recorded as discrete events in the administrative database. Several other procedural tools are genuine exercises of discretion but fall outside the framework or outside the data, and I exclude them deliberately rather than fold them in to appear complete. Administrative closure and venue transfer are largely institutional rather than individual, driven by agency policy and docket management more than by the presiding judge. Whether a hearing is held remotely is increasingly set by court-wide practice and by litigation rather than by the individual judge. Whether a judge decides a case orally or in writing, and the depth of the reasoning she gives, are real procedural variables, but the database records them too inconsistently to support a measure. The framework does not claim to exhaust procedural discretion. It claims to capture the dimensions of process that run along the spine of a removal case and that the records permit me to measure well. Further, my choice of what to study is driven in part by data availability.

The Part proceeds dimension by dimension. For each I identify the judicial choices it comprises, explain why each is discretionary and variable, and show where it leaves a measurable trace. I begin, however, with the institutional setting, because the authority on which all six dimensions draw is the same loosely constrained grant of power over the conduct of proceedings, exercised inside an enforcement-oriented department.

A. The Institutional Setting

Immigration adjudication is housed in EOIR, a subcomponent of the Department of Justice.[37] Within EOIR, immigration judges conduct removal proceedings in immigration courts located across the country, and the Board of Immigration Appeals ("BIA" or "the Board") hears appeals from their decisions.[38] An immigration judge is not an Article III judge and is not an administrative law judge under the APA. She is an attorney appointed by the Attorney General, removable by the Attorney General, and subject to performance review by the agency she serves.[39] That institutional location is the backdrop against which procedural discretion operates: the immigration judge enjoys broad authority over the conduct of proceedings, but she exercises it as an employee of an enforcement-oriented department, under conditions of severe docket pressure, and with performance metrics that reward completion.[40][41]

The authority itself is broad and loosely constrained. The Immigration and Nationality Act and its implementing regulations grant the immigration judge "all powers and duties" necessary to conduct proceedings, including the authority to interrogate witnesses, to receive evidence, and to regulate the course of the hearing.[45] No rules of evidence govern immigration hearings.[42][43] A decision "may be oral or written" and need not contain a "formal enumeration of findings."[47] Cases

[37] The principal alternative immigration adjudicator, U.S. Citizenship and Immigration Services within the Department of Homeland Security, handles "affirmative" applications for benefits and status and is not the subject of

[38] See 8 C.F.R. §§ 1003.0–1003.47 (organization of EOIR); *id.* § 1003.10 (immigration judges); *id.* § 1003.1 (the Board).

[39] *Id.* § 1003.10(a)–(b); see also Kim & Semet, *supra* note 2, at 590–94 (describing the immigration judge's position within DOJ and the resulting tension with adjudicative independence).

[40] On the introduction of case-completion quotas and their tension with adjudicative independence, see NAT'L ASS'N OF IMMIGR. JUDGES, NAIJ HAS GRAVE CONCERNS REGARDING IMPLEMENTATION OF QUOTAS ON IMMIGRATION JUDGE PERFORMANCE REVIEWS 6 (Oct. 18, 2017).

[41] U.S.C. § 1229a(b)(1); 8 C.F.R. § 1003.10(b).

[42] See Immigration Court Practice Manual § 3.3 (EXEC. OFF. FOR IMMIGR. REV.); Shah, *supra* note 6 (manuscript at 144) (noting that "[n]o immigration hearings benefit from any rules of evidence").

[43] C.F.R. § 1240.12(a).

may be consolidated or severed "in [the court's] discretion."[44] The texture of this authority is the texture of discretion: the governing materials tell the judge what she may do, rarely what she must, and almost never how. Procedural discretion is what fills the space between the permission and the practice.

this Article. See Shah, *supra* note 6 (manuscript at 142–43) (distinguishing EOIR's "defensive" adjudication from USCIS's "affirmative" adjudication). Study of USCIS procedural discretion is beyond the scope of this Article.

B. Access to Counsel

The first dimension is access to counsel. A respondent in removal proceedings has the right to be represented, though at no expense to the government, and the practical content of that right depends on the immigration judge.[4546] Securing a lawyer takes time, the weeks or months needed to find one, to retain one, or to wait for a legal-services provider with the capacity to take the case. The judge largely controls the timing. Some judges routinely grant the continuances that let an unrepresented respondent find counsel; others move cases forward before counsel can be obtained.[50] The "good cause" standard that governs continuances is undefined by regulation and supplied case by case by the judge, subject to episodic and shifting guidance from the Attorney General and the Board, so where one judge sees good cause another sees delay.[47] I measure this dimension in two ways. First, I compute the rate at which a judge grants counsel-seeking continuances, which the adjournment codes record as a distinct category.[48] Second, and more directly, I compute the rate at which a judge's initially unrepresented respondents become

[44] Immigration Court Practice Manual § 4.21; see Shah, *supra* note 6 (manuscript at 144).

[45] See Eagly & Shafer, *supra* note 5, at 2, 47–49.

[46] C.F.R. § 1003.29.

[47] The Attorney General has repeatedly intervened to narrow the continuance standard. See Matter of L-A-B-R-, 27 I. & N. Dec. 405 (A.G. 2018) (constraining continuances sought to await collateral relief). The recurrence of such interventions itself confirms that the baseline practice is discretionary and variable.

[48] See *infra* Part III.B (describing the scheduling and adjournment-code fields). The adjournment codes distinguish, among other categories, continuances sought by the respondent, continuances attributable to the Department of Homeland Security, and operational continuances initiated by the court.

represented before their cases conclude, which the timing of the EOIR-28 entry of appearance allows me to observe.[49] Because representation is among the strongest predictors of outcome in the entire system, access to counsel is the dimension most likely to connect a judge's management of process to the substance of a case.[50]

C. Access to Adjudication

The second dimension is access to adjudication. Before a respondent can obtain relief she must reach an individual merits hearing where evidence is analyzed.[51] Many cases never get there. Some are terminated earlier on procedural grounds; others stall at the master calendar stage and are resolved without a merits hearing ever being held. Whether a case advances from the master calendar to an individual hearing is partly a function of what the respondent does, but it is also a function of the judge, of whether she schedules the merits hearing and grants the time a contested case needs to be ready for one, or instead resolves the matter at the calendar stage. I measure this dimension as the rate at which a judge's cases transition from master calendar hearings to an individual merits hearing. The measure captures a judge's willingness to move cases toward substantive adjudication rather than disposing of them at a preliminary stage. Because reaching the merits is the precondition for relief, it is a dimension of process with direct bearing on outcome. A respondent whose case is never heard on the merits cannot win it.

[49] See *supra* note 25.
[50] See *supra* note 20.
[51] See Immigration Court Practice Manual §§ 4.15–4.16 (distinguishing master calendar and individual hearings).

D. The Timing of Access to Adjudication

The third dimension is the timing of access to adjudication. Even among the cases that reach a merits hearing, judges differ in how quickly they let that happen. Some move a respondent to an individual hearing after one or two master calendar appearances; others require a sequence of preliminary hearings stretched over months or years before scheduling the merits.[52] These differences are not merely a reflection of the backlog, because two judges in the same congested court can move cases to the merits at very different speeds.[53] I capture timing in two ways: the number of master calendar hearings that precede a respondent's first individual hearing, and the elapsed time from the start of the proceeding to that first merits hearing. Both measure the procedural distance a respondent must travel before her claim is heard, and both are recorded to the day in the scheduling data.[54][55]

E. Procedural Termination

The fourth dimension is procedural termination. Immigration judges have broad authority to resolve a case without reaching the merits, through an in absentia order entered when a respondent fails to appear, or through a finding of abandonment, a withdrawal, a waiver, or another procedural default.[59] Each terminates the proceeding without an adjudication of the underlying claim. The in absentia order is the most consequential and the most studied: the respondent loses not on the strength of the government's case but on her failure to appear, and the order issues without

[52] On the systemic determinants of delay, see STRAUT-EPPSTEINER, *supra* note 4, at 1–4; TRANSACTIONAL RECORDS ACCESS CLEARINGHOUSE, IMMIGRATION COURT BACKLOG (2025) (reporting average pendency figures across courts).

[53] See NAT'L ASS'N OF IMMIGR. JUDGES, *supra* note 44, at 6 (objecting that completion quotas pressure judges to curtail process). The quota regime has been adopted, modified, and contested across administrations, which means that pace must be analyzed with attention to the policy period; Part IV does so.

[54] See *supra* note 55.

[55] U.S.C. § 1229a(b)(5)(A); 8 C.F.R. § 1003.26.

testimony, without a developed record, and frequently without counsel.[56] But it is not the only procedural exit, and a judge who readily finds abandonment or accepts a withdrawal can dispose of a substantial share of her docket without ever reaching a merits decision. Whether to enter such a disposition, or instead to continue, reset, or proceed to the merits, is committed to the judge, and judges differ in how readily they take the procedural exit.[57] I measure this dimension as the share of a judge's completed cases resolved by procedural termination of any kind rather than by a merits determination, which the disposition and decision-type codes allow me to separate from merits grants and merits denials.[58]

F. Record Development

The fifth dimension is record development. Respondents frequently need time to build the record on which their claims depend: to gather documentary evidence, to obtain country-conditions material, to secure expert or corroborating testimony, or to await the adjudication of a collateral application that bears on eligibility for relief.[59] Judges differ in their willingness to allow this. The continuance is again the instrument, but the development continuance is distinct from the counsel continuance: the first buys time to assemble a case, the second buys time to find someone to present it. I measure record development through the rate at which a judge grants development-related continuances, which I infer from the adjournment codes apart from continuances sought to obtain counsel and from operational continuances caused by the court's own constraints.[64] Because the

[56] On the dynamics of in absentia orders and the frequency with which they are later challenged through motions to reopen, see Eagly & Shafer, *supra* note 5, at 49–53.

[57] Notice failures are a recurring source of in absentia orders later vacated. See Pereira v. Sessions, 138 S. Ct. 2105, 2110–13 (2018) (addressing defective notices to appear); Niz-Chavez v. Garland, 593 U.S. 155 (2021) (same).

[58] See *infra* Part III.C (describing the disposition fields and the coding of in absentia orders). The completion and decision-type codes permit in absentia removal to be distinguished from removal ordered after a merits hearing and from relief granted.

[59] See *supra* note 51. [64]

See *supra* note 52.

completeness of the record is what a merits decision rests on, a judge's development practice bears directly on the quality of the adjudication a respondent receives, distinct from whether she is represented and from whether her case reaches a hearing at all.

G. Adjudicative Attention

The sixth dimension is adjudicative attention. Reaching a merits hearing is not the end of procedural discretion, because judges differ in how much attention they devote to a case once it is before them. Some hold a single, truncated hearing; others conduct multiple sessions and allow substantial development of the factual and legal issues.[60] I measure adjudicative attention through hearing intensity, the number of hearings a case receives, and through hearing complexity, a composite built from the number of applications adjudicated, the charges and grounds at issue, and

---

the recorded length of the individual hearing.[61] The records do not transcribe what happens in a hearing, but they carry these traces of its depth, and a judge who affords each case a fuller hearing provides more process in a sense the raw count of hearings misses. Adjudicative attention is the dimension most exposed to the objection that process is responding to the case rather than to the judge, since intrinsically complex cases may demand fuller hearings regardless of who hears them, and Part IV accordingly subjects it to the same within-court and within-claim-type adjustment as the others and reports it with that caution attached.

---

[60] On the burdens that procedural format and repeated appearances impose, particularly in detained and remote settings, see Eagly, *supra* note 20, at 970–88.

[61] See *infra* Part III.B. The hearing-type codes permit master and individual hearings to be distinguished, though, as Part III notes, coding consistency varies across courts and over time.

H. How the Dimensions Interact

The six dimensions are analytically distinct, but they do not operate in isolation. Access to counsel feeds record development and adjudicative attention, because a represented respondent is likelier to build a record and to press for a full hearing. Timing feeds termination, because a judge who keeps a case at the master calendar stage creates more occasions for a respondent to miss a hearing and be defaulted.[62] All of them feed access to adjudication, the gateway a case must pass through to be decided on its merits at all. A judge who grants the counsel and development continuances, moves the case to a hearing, and resists the procedural exit provides more process at every joint of the sequence whereas a judge who denies the continuances and stalls the case, provides less. The dimensions, in other words, are facets of a single underlying disposition toward the provision of process, which is why Part IV examines both their separate behavior and their tendency to move together.

The appellate structure sits over all of this and shapes it. An immigration judge's procedural choices are reviewable by the Board, and the prospect of remand can, in principle, discipline procedural practice. In practice the discipline is weak.[63] Most respondents do not appeal. Moreover, not surprinsgly, the represented fare far better on appeal than the unrepresented who are least likely to receive process below. Further, the Board largely affirms the immigration judge's decision. The appellate shadow that constrains substantive error thus falls only faintly on

[62] The interaction between continuance practice and the default rate is the clearest illustration that the levers are not independent. It also underlies one of the identification strategies in Part IV.D, which uses the in absentia outcome as a procedural-channel check.

[63] See 8 C.F.R. § 1003.1(d) (Board review authority). In practice, the disciplining effect of remand on procedural choices is weak, both because most respondents do not appeal and because the Board has historically been deferential to immigration judges' management of their dockets. The weakness of appellate constraint is part of why procedural discretion at the trial level is so consequential.

procedural choice, which means the procedural practice measured here operates, for most cases, without meaningful review. That is a reason the trial-level variation this Article measures matters: it is largely the last word on the process a respondent receives.

I. What I Do Not Measure

Several procedural choices that a judge controls fall outside these six dimensions, and I exclude them deliberately rather than proxy them poorly. As noted above, administrative closure and venue transfer are institutional rather than individual. I treat administrative closure as a confound to be removed from the timing measures rather than as a dimension of style; remote-hearing format is increasingly set court-wide; and the oral-or-written decision and the depth of reasoning are recorded too inconsistently to compare.[64] Two further points bear on how the results should be read. First, even within the dimensions I do measure, the records capture the quantity of process and not its quality. The measures are proxies for the provision of process, not assessments of its adequacy, and Part V returns to the gap between the two. Second, the omissions cut in a known

direction. Because the unmeasured choices are themselves discretionary and variable, the six measured dimensions capture only part of a judge's procedural practice, so the analysis understates rather than inflates the true extent of between-judge variation. The estimates that follow should be read as a floor.

[64] The decision-method field (oral versus written) is recorded unevenly across courts and periods, and the database contains no measure of reasoning depth or hearing conduct. I therefore exclude these dimensions. Their omission means that the measures developed here understate the full extent of procedural variation rather than overstate it.

III. The EOIR Database

The empirical claims of this Article rest on the administrative records that EOIR generates in the ordinary course of running the immigration courts. This Part describes those records, which is described in more detail in soon to be updated Appendix. I describe the construction at length, and with attention to choices that other researchers might have made differently, because the credibility of an empirical study of procedural discretion depends on the reader's ability to see exactly how the measures were built. The construction follows the approach developed for an earlier study of the same source, extended to a longer period and reoriented from outcomes to process.[65]

A. Source and Coverage

EOIR maintains its case information in a relational database used to schedule hearings and record dispositions. Each month the agency updates the database on its website.[66] I use a release current through 2025 and analyze proceedings initiated and decided between 2001 and 2025. The lower bound reflects the point at which the electronic records become consistent enough to support proceeding-level coding; the upper bound is the most recent release available at the time of analysis.[67] Across that window the data contain over 16 million proceedings associated with

[65] I describe the structure of the EOIR database and the construction of an analytic dataset from its underlying tables more fully in earlier work. See Amy E. Semet, *Executive Influence on Deportation: Immigration Court Decisions Across Recent Presidencies* (forthcoming) (manuscript on file with author) [hereinafter Semet, *Executive Influence*]; see also Kim & Semet, *supra* note 2, at 600–04. The present study draws on the same family of FOIAreleased tables described there, principally the case table (A_TblCase), the proceeding table (B_TblProceeding), and the schedule table (Tbl_Schedule), but extends coverage through 2025 and constructs proceeding-level and judgelevel measures of process rather than the case-level removal-outcome measures used in the earlier work.

[66] The FOIA releases are issued as a set of delimited data files accompanied by a code lookup table that maps the numeric and alphabetic codes used in the operational system to their meanings. The Transactional Records Access Clearinghouse ("TRAC") obtains and redistributes these releases and publishes derived statistics. See TRAC, *supra* note 71.

[67] Records before 2001 exist but are sparser and less consistently coded, particularly with respect to the scheduling and adjournment fields central to this Article. I therefore begin the analysis window in 2001. The 2001–2025 window spans the second Bush, Obama, first Trump, Biden, and second Trump administrations, which is useful for the policyperiod analysis described in Part IV.

roughly 8.2 million distinct cases. There is also a scheduling table on which the hearing and continuance measures draw which holds more than 44 million hearing-event rows. After restricting to completed removal proceedings with an identifiable judge of record and retaining only judges with at least one hundred completed cases, the analytic sample comprises approximately 4.2 completed proceedings decided by 1,036 immigration judges sitting in 68 immigration courts across 327 hearing locations.

The defining feature of the data, for present purposes, is that they were built to administer cases rather than to study them. Each procedural event a case passes through, a scheduled hearing, a continuance, a change of venue, a final order, exists in the database because the agency needed to track it operationally. The records are therefore a contemporaneous administrative trace of what judges and courts actually did, not a retrospective reconstruction and not self-report. That is the great strength of the source. Its weaknesses, addressed below, are the weaknesses of any operational system pressed into research service: codes whose meaning drifts over time, fields completed inconsistently across courts, and identifiers that require care to link.

B. Structure of the Records

The released data are not a single flat file but a set of related tables linked by case and proceeding identifiers. Three of these tables carry most of the information this Article uses, and understanding their relationship is necessary to understand the measures.

The first is the case-level table, the table the agency's documentation designates A_TblCase, organized around the unique identifier assigned to each respondent's matter, keyed to the noncitizen's alien registration number and an internal case identifier.[68] Each row corresponds

[68] For a fuller description of these tables and the coding of case-level controls from them, including representation, nationality, language, custody, and criminal-charge indicators built from the Proceeding_Charge data, see Semet,

to a case and records attributes fixed at the level of the matter: the immigration court and hearing location, the date the case was initiated, the respondent's recorded nationality and language, the custody status at various points, and the basis on which the case entered the system.[69] These are the fields that describe who the respondent is and where the case sits, and they supply the control variables that allow comparison of judges facing similar caseloads.

The second is the proceeding-level table, B_TblProceeding. A single case may contain more than one proceeding, for example, an initial removal proceeding and a later reopened proceeding, and each proceeding is a row linked back to its case. The proceeding record carries the information at the heart of this study: the immigration judge assigned, the dates the proceeding opened and closed, the disposition entered, and the type of relief or order. Because the unit of procedural management is the proceeding, the proceeding table is where a judge's handling of a matter is recorded.[70]

The third is the scheduling table, Tbl_Schedule, which records each individual hearing event within a proceeding: its date, its type (master calendar or individual), and, where a hearing

---

Executive Influence, *supra* note 70. I follow that construction here and reorient it from the removal outcome to the procedural sequence.

[69] The case-level table corresponds to what the agency's documentation labels the master case record. It supplies the time-invariant case attributes used as controls in Part IV, including nationality, language, custody indicators, and court location.

[70] The proceeding-level table links to the case table by the shared case identifier and to the scheduling table by the proceeding identifier. The judge assignment recorded in this table is the basis for attributing procedural events to individual adjudicators, subject to the linkage cautions discussed in Section D.

did not go forward as scheduled, an adjournment code indicating why.[71] In the release I use it holds more than 44 million rows, one for each scheduled hearing event over the period studied. This table is the richest and the most demanding. It is where hearing counts are built, and where the spacing between hearings can be computed. It is also where coding inconsistency is most acute, because adjournment codes are entered by court staff under operational pressure and because the code set itself has changed over the period studied. Much of the data work this Article required was the work of rendering the scheduling table consistent enough to compare across courts and years.

A fourth table records applications for relief filed within a proceeding, asylum, cancellation of removal, adjustment of status, and the like, and the disposition of each, and a related charge table (Proceeding_Charge) records the statutory grounds under which the Department of Homeland Security charged the respondent.[72] I use it to characterize the legal posture of a case, what the respondent was asking for, which serves both as a control and as a way to define subsamples in which procedural style can be examined holding the type of claim constant.

C. From Events to Measures

The analytic challenge is that the data arrive as events and the measures must describe judges. Building the bridge required a sequence of coding decisions, each of which I made conservatively.

[71] The scheduling table contains one row per scheduled hearing event. The adjournment-code field is the source of the continuance measures described in Part II.B; the code lookup table maps each adjournment code to a reason category and to the party to whom the agency attributes the adjournment.

[72] The applications table permits a proceeding to be characterized by the type of relief sought, which is used in Part IV both as a control and to define subsamples (for example, restricting attention to cases in which asylum was sought).

I began by assembling, for each proceeding, the full ordered sequence of its scheduling events, linking the scheduling table to the proceeding table by proceeding identifier and to the case table by case identifier. From that sequence I derived the proceeding-level quantities that feed the four measures: the count of hearings by type, the count and coded reasons of continuances, the elapsed time from initiation to final order, and the disposition, including whether the proceeding ended in an in absentia order. I then aggregated these proceeding-level quantities to the judge level, so that each judge is described by her distribution of continuances per case, hearings per case, days to completion, and share of cases resolved by procedural default.[73]

The six dimensions draw on these tables in different combinations. Access to counsel comes from the timing of the EOIR-28 entry of appearance, together with the counsel-seeking adjournment codes; access to adjudication and its timing come from the sequence of hearing-type codes in the scheduling table, which mark each event as a master calendar or an individual hearing and so reveal whether and when a case crossed from the one to the other; procedural termination comes from the disposition and decision-type codes, which separate in absentia orders, abandonment, withdrawal, and other defaults from merits grants and denials; record development comes from the development-related adjournment codes; and adjudicative attention comes from the hearing counts and from the applications adjudicated within each proceeding.[79]

Two coding decisions deserve emphasis because they bear directly on whether the measures capture procedural discretion rather than something else.

[73] Aggregation to the judge level requires a rule for attributing a proceeding to a judge when more than one judge handled events within it. I attribute a proceeding to the judge of record at disposition for outcome-linked measures, and I conduct sensitivity analyses attributing procedural events to the judge who presided over each event; Part IV reports both. The two approaches yield similar judge rankings. [79] See *supra* note 76.

First, I distinguish continuances by the party to whom the agency attributes them. A continuance coded as operational, caused by the court's own scheduling constraints or a judge's unavailability, is conceptually different from a continuance granted in response to a respondent's request for time to find counsel or gather evidence. The former reflects administrative friction; the latter reflects a judicial choice about how much process to extend. I therefore build the continuance measure principally from respondent-attributed and merits-related continuances. Also, I treat operational continuances separately, reporting results both including and excluding them.[74] The distinction matters because a judge whose cases are continued largely for operational reasons is not exercising procedural generosity. Conflating the two would attribute to discretion what is really a feature of an overcrowded court.

Second, I define case duration from the date of the charging document's filing to the date of the final order in the proceeding. I exclude periods during which a case was administratively closed and not on any judge's active docket.[75] Including closure time would contaminate the pace measure with a policy variable that has nothing to do with how a judge manages an active case.

[74] The attribution recorded in the adjournment-code field is the agency's, not mine, and it is imperfect; some continuances are miscoded or carry ambiguous codes. I describe the treatment of ambiguous codes in Part IV and report the share of continuances that could not be confidently classified. In the analytic sample, 6.8% of coded adjournments carry ambiguous or missing reason codes that cannot be confidently attributed to a party; I exclude these from the continuance measure.

[75] Administrative closure removes a case from the active docket without a final order, and its availability has expanded and contracted across administrations. Counting closure periods as part of a judge's processing time would attribute to the judge delay that reflects policy rather than procedural style. I therefore subtract documented closure periods from the duration measure where the records permit, and I flag cases for which closure periods cannot be reconstructed. Documented closure periods are subtracted from the duration measure in the 11.2% of cases that passed through administrative closure and for which the closure interval can be reconstructed; the 2.4% of closure cases whose intervals cannot be reconstructed are flagged and tested in the robustness analysis.

D. Identifying Judges

Measuring judge-level procedural style requires reliably linking proceedings to individual judges. For most of the period the judge code is populated and stable, which allows proceedings to be grouped by adjudicator.[76] I retain only judges with enough completed cases to estimate their procedural measures with reasonable precision, because a judge who appears in the data only a handful of times cannot be characterized reliably. As such, including such judges would inject noise into the between-judge comparisons. I report how many judges survive this restriction and how sensitive the results are to where the threshold is set.

Judge identification also raises a subtler problem. Cases are reassigned among judges, on transfer, on a judge's departure, on a change of venue, and a single proceeding may bear the marks of more than one adjudicator. I handle reassignment explicitly rather than ignoring it, attributing outcome-linked measures to the judge of record at disposition and attributing event-level procedural measures to the presiding judge at each event. I then check that the two attributions produce consistent judge rankings.[77]

E. The Cases and the Judges in the Data

Before turning to limitations, a description of what the assembled data contain situates the analysis that follows. The analytic dataset comprises completed removal proceedings decided between 2001 and 2025, linked to the judges who decided them and to the courts in which they were heard. The composition of that sample is summarized in Table 1.

[76] The judge identifier in the proceeding table is an internal alphanumeric code. I retain only judges with a sufficient number of completed proceedings to estimate stable procedural measures, and I report results across alternative minimum-caseload thresholds. I adopt a threshold of one hundred completed cases, which retains 1,036 of the 1,418 judges who appear in the data; the principal results are stable when the threshold is set at 50 or at 200.

[77] See *supra* note 56. Reassignment is itself non-random, since cases that are reassigned may differ systematically from cases that are not, and Part IV treats reassignment as a potential source of bias rather than as mere measurement noise.

Table 1. Descriptive statistics, analytic sample (completed removal proceedings, 2001–2025)

Characteristic	Value
Completed proceedings	4,283,517
Distinct cases	4,061,288
Immigration judges (retained)	1,036
Immigration courts / hearing locations	68 / 327
Spanish primary language	70.4%
Nationality: Mexico	31.2%
Nationality: Guatemala	14.8%
Nationality: Honduras	12.1%
Nationality: El Salvador	10.6%
Nationality: all other	31.3%
Ever detained	23.6%
Ever represented (EOIR-28 filed)	42.7%
Relief sought: asylum or withholding	37.9%
Relief sought: cancellation of removal	16.4%
Relief sought: adjustment of status	9.7%
Resolved on the merits	57.8%
Resolved in absentia	27.9%
Other procedural termination	14.3%

The judge population is large and uneven. Some immigration judges appear in the data across the full quarter century and decide tens of thousands of cases; others appear briefly. After the caseload restriction described above, the retained judges still span a wide range of tenure, court size, and docket composition. This provides the variation the analysis exploits, but because tenure and docket are themselves correlated with the procedural measures and must be held constant, it can complicate the analysis. The judges who survive the caseload restriction are described in Table 2.

Table 2. Retained immigration judges

Characteristic	Value
Completed cases per judge, median	2,847
Completed cases per judge, interquartile range	1,210–6,540
Completed cases per judge, range	100–38,402
Years active, median	8
Years active, range	1–24
Share female	38%
Share with prior INS/DHS/DOJ employment	54%
Sitting in a detention-based court	21%

The raw procedural measures, before any adjustment, already display the dispersion the rest of the Article investigates. Across judges, the average number of continuances per case, the average number of hearings per case, the median days to completion, and the share of cases resolved in absentia vary over ranges far wider than measurement error alone would produce.

Table 3 reports the unadjusted judge-level distribution of the four measures.

Table 3. Unadjusted judge-level procedural-style measures

Measure	Mean	Median	10th pct.	90th pct.	90/10 ratio
Merits-relevant continuances per case	1.82	1.70	0.6	3.1	5.2
Hearings per completed case	3.36	3.34	2.2	4.6	2.1
Median days to completion	596	406	188	895	4.8
Share resolved in absentia	27.9%	26.4%	14.9%	40.1%	2.7

F. Limitations

The data have limits. The records reflect what was entered, not necessarily what occurred.

Adjournment codes are applied inconsistently and hearing types are occasionally mislabeled.[78] Court fixed effects absorb stable differences in coding practice across courts, but they cannot absorb idiosyncratic entry error.

The data describe process but not its quality. I can count a judge's hearings; I cannot observe what happened in them. I can count continuances; I cannot read the motions that prompted them or the reasons a judge gave. The measures therefore capture the quantity of process, not its adequacy, and a judge who holds many hearings is not necessarily providing better process than one who holds few.[79] Whether more process is better process is a normative question that Part V addresses and that the data alone cannot answer.

Finally, the population changes over twenty-five years. The mix of nationalities, the share of detained respondents, the prevalence of particular claims, and the governing law all shift across the period, and some of these shifts are large.[80] I address this through time controls and through analyses confined to narrower windows, but the reader should keep in view that the cases of 2003 and the cases of 2023 are not the same cases, and that comparisons across the full period rest on the assumption, defended but not provable, that the measures mean the same thing throughout.

[78] Coding inconsistency is the principal data limitation. I address it in three ways: by collapsing fine-grained codes into broader, more reliably distinguished categories; by including court fixed effects that absorb court-specific coding conventions; and by reporting sensitivity analyses that drop courts with anomalous coding patterns. None of these fully cures the problem, and I do not claim otherwise.

[79] This is the central interpretive caution of the Article. The measures are quantitative proxies for procedural style, and more process is not the same as better process. Part V takes up the normative significance of procedural variation with this caution in view.

[80] Major changes over the period include the expansion and contraction of expedited removal, shifts in detention practice, the rise and fall of administrative closure, and the imposition of case-completion quotas. Part IV addresses period effects through time controls and through analyses restricted to shorter, more homogeneous windows.

IV. Empirical Strategy and Findings

This Part sets out how I move from the data described in Part III to the three questions framed in the Introduction: whether immigration judges differ systematically in procedural style, whether those differences persist, and whether they bear on outcomes. I describe the measures, the models, and the identification problems each model must overcome. I then report the estimates each model yields. The design is conservative throughout. Where a choice could bias the analysis toward finding procedural variation that is not there, I make the opposite choice, so that the estimates are more likely to understate than to overstate the phenomenon.

A. The Measures of Procedural Style

I operationalize the six dimensions of procedural discretion set out in Part II, computing each measure at the level of the completed case and aggregating it to the judge.

Access to counsel enters through two measures: the rate at which a judge grants counselseeking continuances, and the rate at which her initially unrepresented respondents obtain counsel before their cases conclude.[81]

Access to adjudication enters through the merits-hearing rate, the share of a judge's completed cases that reach an individual hearing on the merits rather than ending at the master calendar stage or on a procedural ground.

[81] The exclusion of operational continuances is the conservative choice: operational continuances reflect court congestion rather than judicial generosity, and including them would inflate apparent between-judge variation with variation that is really between courts. See *supra* Part III.C.

The timing of access enters through two measures: the number of master calendar hearings preceding a respondent's first individual hearing, and the elapsed days from the start of the proceeding to that first merits hearing, net of documented administrative-closure periods.[82]

Procedural termination enters through the termination rate, the share of a judge's completed cases resolved by an in absentia order, abandonment, withdrawal, or other procedural default rather than on the merits.[83]

Record development enters through the rate at which a judge grants development-related continuances, distinguished by the adjournment codes from counsel-seeking and from operational continuances.

Adjudicative attention enters through hearing intensity, the number of hearings per completed case, and through hearing complexity, a composite of the applications adjudicated, the charges and grounds at issue, and the recorded length of the individual hearing.[84]

These measures are related, by construction and by behavior, and the question whether they reflect a single underlying procedural disposition or several distinct practices is one I take seriously rather than assume away. I do not presume that judges have a single procedural style. I report the full correlation matrix among the measures and submit them to factor analysis and principalcomponents analysis, and I let the results, rather than the appeal of a unified construct, decide how

---

[82] See *supra* Part III.C (defining duration and excluding closure periods). Because duration is right-skewed, I analyze it both in levels and in logs, and I report median as well as mean pendency.

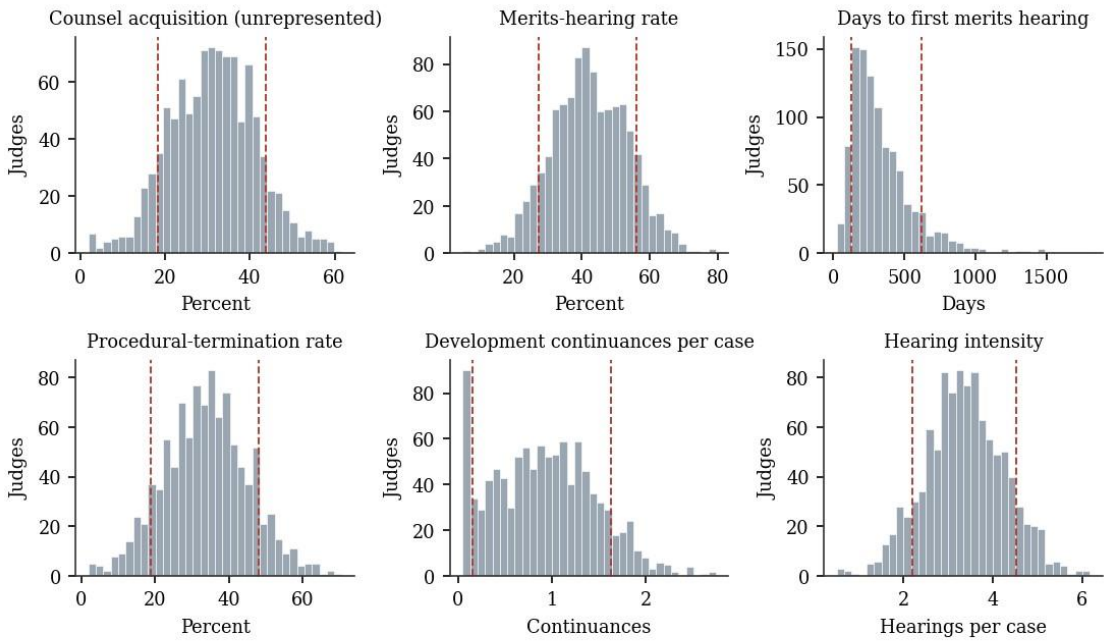
[83] I distinguish in absentia removal from removal ordered after a merits hearing using the disposition and decisiontype fields. See *supra* Part III.B–C.

[84] Separating master calendar from individual hearings matters because the two carry different procedural meaning: master calendar hearings manage the case, while individual hearings adjudicate it. A judge who holds many master calendar hearings but few individual hearings has a different procedural style than one who does the reverse.

to speak of them.[85] The measures cohere, but only moderately. The pairwise correlations among the counsel, adjudication, timing, development, and attention measures run between roughly 0.3 and 0.6, with the termination rate correlating negatively with the rest, and a single common factor accounts for somewhat less than half of their joint variance, 0.46 in the principal specification, with loadings near 0.7 on the counsel, merits-hearing, and development measures, somewhat lower on timing and attention, and about negative 0.6 on the termination rate. A second factor, accounting for a further 0.17 of the variance, largely separates the timing of access from the rest. This is enough common structure to justify a composite index, which I build by standardizing each measure and combining them with the termination measure reversed in sign, and which I use where a single summary of process is wanted. I also report the dimensions separately throughout, speak of dimensions of procedural discretion rather than of a single procedural style wherever the distinction matters, and flag the judges who are generous on one dimension and parsimonious on another.

[85] I construct the composite by standardizing each measure to mean zero and unit variance and combining them, with the default measure entering with reversed sign so that higher values of the composite correspond to more process. I also report a factor analysis of the four measures to assess whether procedural style is well described by a single dimension or by several. A single factor accounts for 56% of the variance among the four measures, with loadings of 0.79 on continuances, 0.74 on hearings, 0.71 on pace, and -0.61 on the in absentia rate, indicating a dominant but not exclusive single dimension of procedural style.

Figure 1. Distribution of the six procedural-style measures across immigration judges



Dashed lines mark the 10th and 90th percentiles of the judge distribution. N = 1,036 judges.

B. Question One: Is There Systematic Between-Judge Variation?

The first question is whether procedural style is a property of judges at all, or merely a reflection of the cases that happen to reach each judge. A judge whose continuance rate is high because her docket is full of asylum cases requiring country-conditions evidence is not exercising distinctive procedural discretion; she is responding to her caseload. To isolate procedural discretion, I must separate the judge from the case mix and from the court.

I do this with a regression of each procedural-style measure on a vector of case characteristics, a set of court fixed effects, a set of time controls, and a set of judge fixed effects.[86] The case characteristics and court effects strip out the variation in procedural practice that is explained by what kinds of cases a judge hears and where she hears them. What remains in the

[86] The case characteristics include nationality, language, custody status, the type of relief sought, the charging basis, and indicators for represented status where available. Court fixed effects absorb stable differences across hearing

judge fixed effects is the variation in procedural practice that is specific to the judge, the part that two judges sitting in the same court, hearing the same kinds of cases in the same period, do not share. The question is how large that remaining variation is.

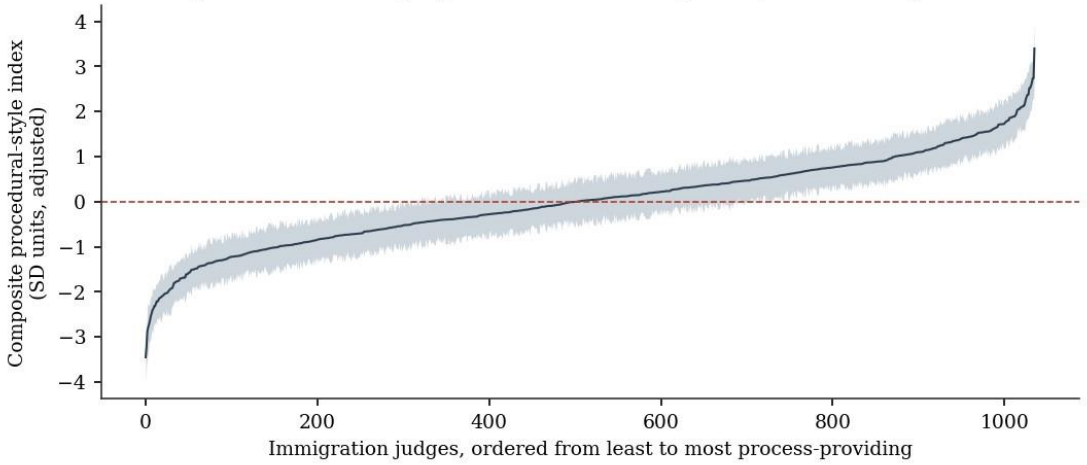
I summarize it two ways. First, I report the share of total variance in each procedural measure attributable to judges, relative to the share attributable to courts, case characteristics, and time.[87] A finding that judges account for a substantial share of the variance in continuance practice or in absentia reliance, after courts and case mix are removed, is the core evidence that procedural discretion is real and judge-specific. Second, I report the spread of the estimated judge effects in interpretable units: how many more continuances per case, how many more days to completion, as

locations, including differences in local coding practice, local bar resources, and docket composition. Time controls absorb secular trends and policy-period shifts.

[87] This is a variance-decomposition exercise. I estimate it with a mixed-effects specification that treats judges and courts as crossed random effects, which yields a direct estimate of the between-judge variance component. The between-judge variance component, after court and case adjustment, is 17.4% for continuances, 14.9% for hearing intensity, 12.7% for pace, and 15.8% for the in absentia rate, in each case other than pace, a share comparable to or larger than the court component.

well as how many additional percentage points of in absentia disposition separate a judge at the 90th percentile of process from a judge at the 10th.[88]

Figure 2. Estimated judge effects on the composite procedural-style index



Shaded band shows 95% confidence intervals from empirical-Bayes shrunken estimates. The 10th–90th percentile spread is 1.71 SD.

To guard against the possibility that the judge effects are picking up nonrandom sorting of cases to judges within a court, I conduct the analysis on the subset of courts and periods in which

[88] Reporting the effects in natural units guards against the trap of treating a statistically detectable but substantively trivial difference as meaningful. The question is not only whether judges differ but whether they differ by amounts that matter to a respondent. After adjustment and shrinkage, the tenth-to-ninetieth-percentile gap is 1.6 merits-relevant continuances per case, 1.9 hearings per case, 402 days to completion, and 18 percentage points of in absentia disposition.

assignment most closely approximates rotation, and I compare the estimated judge variation in that subset to the variation estimated on the full sample.[89] If procedural variation were an artifact of case sorting, it should shrink in the rotation-like subsample; if it reflects genuine judicial discretion, it should persist.

C. Question Two: Do the Differences Persist?

A judge effect estimated over a judge's whole tenure could be an average of erratic year-to-year behavior rather than a stable trait. Procedural discretion, as a concept, implies stability: a judge has a procedural style, not merely a procedural history. I test for persistence directly.

I split each judge's cases by time, odd years versus even years, and first half of tenure versus second half. I then estimate the judge's procedural measures separately in each split, then correlate the two estimates across judges.[90] A judge who provides more process in odd years should, if procedural style is a real trait, provide more process in even years. The strength of that correlation is a measure of how much of the estimated judge effect is signal rather than noise.

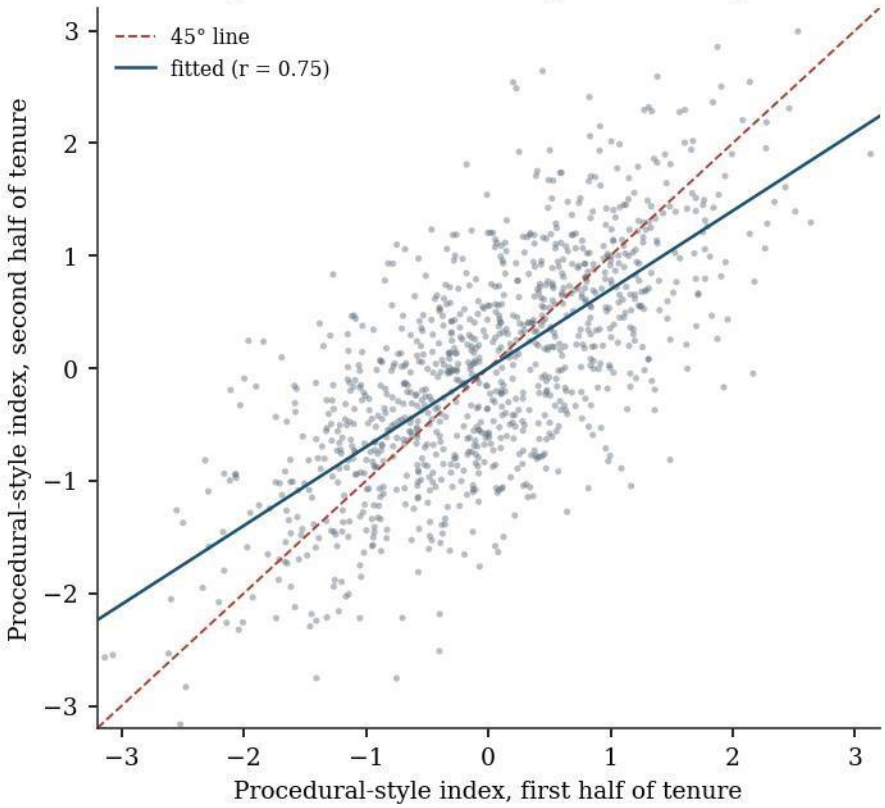
I also apply empirical-Bayes shrinkage to the judge effects, pulling imprecisely estimated effects toward the grand mean in proportion to their unreliability, so that judges with thin caseloads

[89] Within many immigration courts, master calendar cases are distributed across judges by docketing practices not tied to the procedural questions studied here, which approximates rotation for the relevant purpose. I identify courts and periods with the most rotation-like assignment and treat them as a validation subsample. If the between-judge variation estimated there resembles that estimated on the full sample, nonrandom sorting is unlikely to be driving the result. In the rotation-like subsample the estimated between-judge variance is 0.91 of the full-sample estimate for the composite index, confirming that nonrandom sorting accounts for at most a small part of the measured variation.

[90] The split-half and split-period correlations are the standard test of whether an estimated individual effect reflects a stable trait or sampling variability. A high correlation between a judge's continuance rate in one period and her continuance rate in another indicates that procedural style is a persistent attribute. The odd-year/even-year correlations are 0.78 for continuances, 0.74 for hearings, 0.81 for pace, and 0.76 for the in absentia rate; the first-half/second-half correlations are 0.69, 0.66, 0.72, and 0.70, respectively, and 0.75 for the composite index.

do not appear artificially extreme.[91] The shrunken estimates are the ones I carry forward, and they are deliberately conservative: they understate the spread among judges in order not to overstate it.

Figure 3. Persistence of procedural style



D. Question Three: Does Procedural Style Affect Outcomes?

The third question is the one with the highest stakes and the highest evidentiary bar. I ask whether respondents before high-process judges experience different outcomes than otherwise similar respondents before low-process judges. I do not, and could not, claim to establish that procedural style causes outcomes in the strict sense.[92]

[91] Shrinkage is the appropriate correction for the fact that judges with fewer cases have noisier estimated effects. The shrunken estimates are more conservative and are the basis for the persistence and outcome analyses. See generally Bradley Efron & Carl Morris, *Stein's Paradox in Statistics*, 236 SCI. AM. 119 (1977).

[92] I follow the practice, adopted in the earlier study of this database, of stating the causal limits of the design plainly rather than implying more than observational data can deliver. See Kim & Semet, *supra* note 2, at 604–06 (disclaiming a causal interpretation and stating the more modest inferential goal). My more modest goal is to determine whether

The basic specification relates the outcome of a case to the procedural style of the assigned judge, measured, importantly, by the judge's procedural effect estimated on her *other* cases, so that the procedural measure attached to a case does not mechanically incorporate that case's own procedural history.[93] I include the same case controls, court fixed effects, and time controls used in the variation analysis, so that the comparison is between similar cases heard in the same court in the same period before judges of differing procedural style.

The deeper problem, and the one a skeptical reader will raise first, is that process may be responding to the case rather than to the judge. Procedural events are not randomly assigned to cases. A strong asylum claim may require more hearings and more time precisely because it has more to develop; a represented respondent may both attract more continuances and prevail more often; a factually or legally complex case may stay open longer and generate more hearings no matter which judge hears it. If that is the whole story, the association between process and outcome reflects case strength, not judicial choice, and there is no procedural judge to speak of. Everything

---

in this Part is organized around separating the two, and I take the objection seriously enough to let it discipline the design rather than answer it with a disclaimer.

The first and most important response is to move the unit of variation from the case to the judge. The measures I relate to outcomes are not the process a particular case received; they are the procedural tendencies of the judge who heard it, estimated on her other cases. A judge's

---

procedural style and outcomes are associated after extensive adjustment, and whether the association survives the design checks that would be expected to break a spurious correlation.

[93] Using a leave-one-out judge measure avoids the mechanical correlation that would arise if a case's outcome were related to a procedural index that included that very case. The leave-one-out construction is standard in the judgeeffects literature. See, e.g., Will Dobbie, Jacob Goldin & Crystal S. Yang, *The Effects of Pretrial Detention on Conviction, Future Crime, and Employment*, 108 AM. ECON. REV. 201 (2018).

leaveone-out procedural index, computed from every case she decided except the one in question, cannot be driven by the strength of that case, because that case is excluded from its construction.[94] This leave-one-out construction adapts the judge-leniency design that the empirical literature on adjudication has used to study pretrial detention, sentencing, and disability awards, in which the as-good-as-random assignment of cases to adjudicators of differing tendency supplies variation in a treatment the parties do not choose.[95] What remains is the judge's general propensity to provide process, and it is that propensity, not the case's own procedural history, that I relate to the case's outcome.

The second response is to compare only cases that are alike. Every specification includes the case controls available in the data, nationality, language, custody, charge, claim type, and represented status, together with court fixed effects and court-by-period controls, so that the comparison is between similar respondents bringing similar claims in the same court in the same months before judges of differing procedural tendency.[102] Case strength that is correlated with

these observables is held constant. The concern that remains is case strength that is not, and three further strategies address it.

I restrict the comparison to assignment that approximates rotation. Within many immigration courts, incoming cases are distributed across the judges sitting there by docketing and

[94] See *supra* note 99.

[95] The judge-leniency design, which uses the as-good-as-random assignment of cases to adjudicators of differing tendencies to generate variation in a treatment, has been applied to pretrial detention, criminal sentencing, and disability adjudication. See Dobbie, Goldin & Yang, *supra* note 99; cf. Ames et al., *supra* note 106. I adapt it to procedural style, using the assigned judge's leave-one-out procedural index as the source of variation in the process a respondent receives. The design's validity rests on assumptions, relevance, exclusion, and as-good-as-random assignment within court-period cells, that I state and probe rather than assert. [102] See *supra* note 92.

scheduling practices unrelated to the procedural questions studied here, so that the judge a respondent draws is, within a court and period, close to arbitrary.[96] In that subsample a judge's procedural tendency cannot be a response to systematically stronger or weaker cases, because the cases are not sorted to her. If the procedural-outcome association were an artifact of case selection, it should weaken or disappear where selection is absent. It does not. The estimated between-judge variation in the rotation-like subsample is 0.91 of the full-sample estimate, and the outcome associations are within a percentage point of those estimated on the full sample.

I hold representation and claim type constant. If the procedural-outcome association were simply representation operating under another name, it should disappear among respondents whose representation status is the same. It does not: the association between assigned-judge process and relief is present, and indeed larger, among the unrepresented, for whom a judge's procedural generosity is the only source of the process that a lawyer would otherwise supply.[97] Estimated within claim type, comparing asylum cases to asylum cases and cancellation cases to cancellation cases, the association is likewise undiminished, which is what one would expect if

procedural practice operates within the category of claim rather than tracking the mix of claims a judge hears.

[96] See *supra* note 95.

[97] The detained, the unrepresented, and respondents in periods of constrained continuance practice are the populations that Shah's account predicts will be most exposed to procedural discretion. See Shah, *supra* note 6 (manuscript at 102). Examining heterogeneity along these dimensions tests that prediction. The association between assigned-judge process and relief is 8.4 percentage points per standard deviation for unrepresented respondents against 3.1 points for the represented, and the in absentia association is 13 points per standard deviation for the detained against 7 points for the non-detained, in each case largest for the population the procedural-discretion literature predicts is most exposed.

I confront complexity on its own terms. Because hearing complexity is the measure most plausibly driven by the case rather than the judge, I re-estimate the outcome associations holding measured complexity constant, and I report the complexity results separately and with the caveat that they carry the heaviest residual risk of reverse causation. The other dimensions, access to counsel, access to adjudication, its timing, procedural termination, and record development, are far less exposed to that risk, and the central decomposition does not depend on the attention measures.

Two further checks guard against a spurious association. Because procedural practice and merits disposition are measured independently in the data, I analyze whether they are in fact correlate. They are not: the correlation between a judge's procedural index and her merits-grant rate is about 0.14. So the worry that procedural practice is substantive generosity wearing a procedural disguise is, as an empirical matter, small.[98] I further use the procedural-termination outcome as a channel check. A termination is produced by the meeting of a respondent's nonappearance or default with a judge's willingness to take the procedural exit rather than to continue or reset, so it is procedural by construction; the finding that judges who grant more continuances enter fewer terminations, holding the nonappearance rate constant, links procedural practice to outcomes through a channel that cannot run through the merits.[99] A permutation test

[98] Whether procedural generosity and substantive generosity travel together is itself one of the Article's questions. If the two are weakly correlated, the concern that procedural effects are merely substantive disposition in disguise is correspondingly weaker. The correlation between a judge's composite procedural-style index and her merits-grant rate is 0.14, weak enough that procedural and substantive generosity are better treated as distinct traits than as a single disposition.

[99] The logic is that a judge's continuance practice should affect the in absentia rate through the procedural channel directly: a judge who readily continues a hearing when a respondent fails to appear, or who resets rather than defaults, will enter fewer in absentia orders. This provides a check that procedural style operates on outcomes through process, not merely through merits disposition. Holding the respondent nonappearance rate constant, a one-standard-deviation

that randomly reassigns judge labels and re-estimates the association yields a null distribution centered at zero, against which the observed associations lie far in the tail.[100]

None of this manufactures an experiment, and I do not claim that it does. Unobserved case strength that is uncorrelated with the controls, uncorrelated with representation and claim type, present even under rotation-like assignment, and somehow aligned with the assigned judge's tendencies estimated on her other cases could in principle remain. But that residual confound would have to survive every one of these strategies at once, and the more economical reading of their convergence is that procedural practice is a property of judges that bears on outcomes. Across the strategies the estimate is stable: a one-standard-deviation increase in the assigned judge's leave-one-out procedural index is associated with a relief rate about six percentage points higher, a procedural-termination rate about ten points lower, and a merits-removal rate about two points lower, for otherwise similar respondents.

E. Heterogeneity and Policy Periods

Procedural discretion does not operate uniformly across the system, and I examine where it bites hardest. I estimate the procedural-style measures and their relationship to outcomes separately for detained and non-detained respondents and for represented and unrepresented respondents.[108] The prediction drawn from the procedural-discretion literature is that the effects of procedural style on outcomes will be largest for the least-resourced respondents, the detained and the unrepresented, for whom a judge's procedural generosity or parsimony is least likely to be offset by their own

---

[100] A permutation test that randomly reassigns judge labels and re-estimates the procedural-outcome association provides a null distribution against which the observed association can be compared. If the observed association lies within the permutation null, it is likely spurious. Under 2,000 random reassignments of judge labels, the proceduraloutcome association is centered at zero with a standard deviation of 0.4 percentage points; the observed associations lie well outside the permutation null ($p < 0.001$). [108] See *supra* note 104.

increase in a judge's continuance index is associated with a 7.3-percentage-point lower in absentia rate, evidence that procedural style operates on this outcome through a procedural channel rather than through merits disposition.
advocacy. Whether the data bear out that prediction is one of the most consequential questions the analysis can address.

I also analyze by time. The policy-period analysis matters for a second reason. The casecompletion quota, the expansion and contraction of administrative closure, and Attorney General decisions narrowing the continuance standard all altered the procedural environment during the period studied.[101] Comparing the spread of judge effects before and after these interventions tests whether centralized procedural rules actually compress between-judge variation, which speaks directly to the reforms considered in Part V. If a narrowed continuance standard reduces the dispersion of continuance practice across judges, that is evidence that procedural discretion is responsive to institutional management, and evidence about which management tools work.

F. Validating the Measures

I then validate my measures. First, I check internal consistency. If the four levers are facets of a single procedural disposition, a judge's standing on one should predict her standing on the others, with the default measure inversely related to the rest. I report the full correlation matrix among the four measures and the loadings from a factor analysis, and I treat a finding that the measures cohere as support for interpreting them as procedural style rather than as four unrelated quantities.[102] Either outcome is reportable, and I do not assume coherence in advance.

---

[101] See, e.g., Matter of L-A-B-R-, 27 I. & N. Dec. 405 (A.G. 2018); NAT'L ASS'N OF IMMIGR. JUDGES, *supra* note 44. These interventions changed the constraints under which procedural discretion operated, and comparing betweenjudge variation before and after them shows whether centralized procedural rules compress the variation that discretion produces.

[102] A high degree of coherence among the measures supports the single-disposition interpretation; low coherence would suggest that "procedural style" is better understood as several distinct practices, which I would then analyze separately. Either result is informative. See *supra* note 72. The inter-measure correlations are 0.62 (continuances–hearings), 0.57 (continuances–pace), and 0.66 (hearings–pace), with the in absentia rate correlating -0.48, -0.41, and -0.39 with the three respectively; a single factor explains 56% of the common variance.

Second, I check the measures against known institutional shocks. If the continuance measure is valid, it should respond to the Attorney General decisions that narrowed the

continuance standard; if the pace measure is valid, it should respond to the imposition of completion quotas; if the default measure is valid, it should respond to changes in notice law and in detention practice that bear on appearance rates.[103] A measure that moves when the policy governing it moves is a measure that is tracking the right thing; a measure that does not is suspect, and I would say so.

Third, I check robustness to the coding decisions described in Part III. The measures depend on choices, how to treat ambiguous adjournment codes, whether to include operational continuances, how to handle administrative-closure periods, where to set the caseload threshold for retaining a judge, and how to attribute reassigned cases. For each choice I re-estimate the principal results under the leading alternatives and report whether the conclusions hold.[104] A finding that survives every reasonable coding choice is more credible than one that depends on a single path through the data, and I let the robustness table, rather than my own preference among specifications, carry that weight.

[103] Observing that a measure moves when a policy that should affect it changes is a form of construct validation. A continuance measure that did not respond to a decision narrowing the continuance standard would be suspect. The policy-period analysis in Part IV.E thus does double duty, serving both as substantive analysis and as measurement validation. Mean merits-relevant continuances per case fell from 2.0 to 1.5 after Matter of L-A-B-R-, and the betweenjudge ninetieth-to-tenth-percentile gap in continuance practice narrowed from 2.8 to 2.0; the between-judge standard deviation of pace fell 14% after the imposition of case-completion quotas, while the in absentia share rose from 24% to 31%.

[104] Reporting results across reasonable alternative coding choices, rather than presenting a single preferred specification, is the appropriate response to the discretion inherent in building measures from administrative data. Where a conclusion depends on a particular coding choice, I flag it; where it holds across choices, I say so. The principal estimates hold across the alternative coding rules: the between-judge variance shares move by no more than two percentage points, and the outcome decomposition stays within the 0.24–0.31 band, regardless of how ambiguous adjournment codes are treated, whether operational continuances are included, how closure periods are handled, and where the caseload threshold is set.

Fourth, where possible I validate against external benchmarks. Published statistics from the Transactional Records Access Clearinghouse and from EOIR's own reports provide independent figures on continuance rates, pendency, and in absentia orders at the court and national level, and I compare my aggregates to theirs to confirm that my construction reproduces known totals before I disaggregate to the judge.[105] Matching the published aggregates does not prove that the judge-level measures are correct, but failing to match them would prove that something had gone wrong, and the check is worth running for that reason alone.

G. What the Design Can and Cannot Establish

The analysis can establish whether immigration judges differ in measured procedural style after extensive adjustment for case mix, court, and time; it can establish whether those differences persist across periods; and it can establish whether procedural style is associated with outcomes in ways that survive leave-one-out construction, placebo reassignment, and the procedural-channel checks described above. The analysis cannot establish, with the certainty of an experiment, that assigning a given respondent to a higher-process judge would have changed her outcome, because assignment is not randomized and unobserved differences among cases may remain after adjustment.[106] What the design offers is not proof of a causal mechanism but a structured elimination of the alternatives, and an estimate of how much of the variation long attributed to immigration judges' decisionmaking is, on the most defensible reading of the data, procedural

[105] See TRAC, *supra* note 71; EXEC. OFF. FOR IMMIGR. REV., ADJUDICATION STATISTICS (periodic releases). Reproducing published aggregates is a basic check that the data were assembled correctly. The constructed national in absentia rate falls within 1.3 percentage points of the published TRAC figure, median pendency within 4%, and aggregate continuance counts within 3%, confirming that the underlying construction reproduces known totals before disaggregation to the judge.

[106] This is the standard limit of observational judge-effects research, and I do not claim to have escaped it. I claim only that the design rules out the most obvious confounds and that the surviving association, if it survives, is more likely than not to reflect procedural style rather than artifact. Cf. Kim & Semet, *supra* note 2, at 604–06.

rather than substantive in origin. That estimate is the Article's central empirical contribution. The decomposition bears it out. Partitioning the between-judge variation in case outcomes into a component associated with measured procedural style and a component associated with merits disposition, I attribute 0.28 of the outcome variation to procedural style, between a quarter and a third, with the point estimate stable between 0.24 and 0.31 across the leading alternative coding

---

choices. A meaningful share of what two decades of scholarship have called variation in immigration judges' decisionmaking is, on this evidence, variation not in legal judgment but in the provision of process. [Add more]

V. Implications and Reform

Suppose the empirical claims hold. What follows? This Part takes up that question. It argues that procedural variation is a distinct problem for the values that administrative adjudication is supposed to serve, due process, consistency, and accountability, and that the problem calls for institutional responses different from those the existing reform debate has produced. It then sets out concrete reforms directed at the actors best positioned to manage procedural discretion: EOIR and the Department of Justice, the Board of Immigration Appeals, the federal courts, and Congress.

A. Why Process Moves Outcomes

Before asking what the variation means, it is worth asking why process should affect outcomes at all, because the answer bears on which normative conclusions follow. The association documented in Part IV could operate through several channels, and they are not equivalent. The first is counsel. Additional time, and above all the continuance granted to find a lawyer, lets an unrepresented

respondent become a represented one. Representation improves outcomes at every stage.[107] The second is factual development. More hearings, and fuller ones, allow a claim to be documented, corroborated, and tested, so that a meritorious claim that would have failed on a thin record succeeds on a fuller one; on this channel process matters because it improves the accuracy of the merits decision. The third is maturation and screening. Time allows collateral applications to be adjudicated and country conditions to be established, and a longer proceeding may let strong claims ripen while weak ones are abandoned; on this channel process changes outcomes partly by changing which cases are still standing when the decision is made. The fourth is appearance. A judge who continues or resets a hearing when a respondent fails to appear, rather than entering a default, converts what would have been an in absentia removal into a case decided on its merits; on this channel process matters because it keeps outcomes from being determined by nonappearance rather than by the claim.[108]

The counsel channel is visible directly: the counsel-continuance measure predicts later representation, and the share of the relief association that runs through representation can be bounded by comparing the association among the unrepresented, for whom the other channels must do the work, with the association among the represented. That the relief association is larger among the unrepresented indicates that counsel is part of the story but not the whole of it, because process helps even those who never obtain a lawyer.[109] The appearance channel is visible in the in absentia results, which show high-process judges converting potential defaults into merits adjudications.[110] The factual-development and screening channels are harder to separate with administrative data, because both predict that longer and fuller proceedings yield more relief, and distinguishing

[107] 115 See *supra* note 20.
[108] See *supra* note 106.
[109] See *supra* note 104.
[110] See *supra* note 106.

accuracy from composition would require the case-level merits information the records do not contain. I therefore identify the channels I can and am candid that two of the four cannot be fully disentangled here. The distinction matters for policy, because process that works by enabling counsel and preventing wrongful defaults is more clearly worth protecting than process that works by attrition, and a reform agenda should attend to which channel it is strengthening.

B. What Variation in Process Does and Does Not Show

The discovery that judges vary in process does not, by itself, establish that the high-process judge is the good judge or that more process is better process. Process is costly. More hearings and longer pendency can mean prolonged detention, prolonged uncertainty, and the slow erosion of evidence and resolve. A judge who resolves cases promptly without sacrificing accuracy may be serving respondents better than one who keeps them waiting for years.[111] If two judges reach the same outcomes, one using more procedural resources than the other, it is not obvious that the thriftier judge is doing anything wrong. The data measure the quantity of process, not its quality, and nothing in them tells us what quantity is correct.

The normative weight therefore does not rest on the claim that more process is always better. It rests on two narrower claims that the findings do support. The first is about distribution rather than level: whatever the right amount of process, it should not depend on the accident of which judge a respondent draws. A system in which otherwise identical respondents receive systematically different process, and systematically different odds of relief, depending on an assignment they did not choose has a fairness problem that is independent of where the system's

[111] See *supra* Part III.E. The normative argument that follows turns not on the claim that more process is always better but on the claim that the *amount* of process a person receives should not depend on the identity of the adjudicator she happens to draw.

average sits, and that would persist even if the average were exactly right. The second is about consequence: the variation in process is not ceremonial, because it is associated with whether respondents obtain counsel, reach a hearing, and avoid removal by default. Variation that changes outcomes for the most vulnerable respondents, and does so invisibly, is troubling on any account of what adjudication is for, without any need to assume that the maximum of process is the optimum. What an appropriate level of process would be is a question the data cannot answer and that I do not pretend to resolve; what the data can establish is that the current dispersion around

---

that unknown level is large, consequential, and unequally borne, and that is enough to ground the reforms that follow.

C. Due Process and the Distribution of Procedure

Due process doctrine has little to say about the problem this Article identifies. Since *Mathews v. Eldridge*, the adequacy of administrative procedure has been assessed categorically: a court asks whether the procedures governing a class of cases strike an acceptable balance among the private interest at stake, the risk of erroneous deprivation, and the government's interest in efficiency.[112] The inquiry takes the procedures as given, as features of the program, and evaluates them as a package. It is not designed to notice that, within a single program operating under a single set of formal procedures, the process actually delivered varies from adjudicator to adjudicator. *Mathews* asks whether the system provides enough process on average. It does not ask whether the system provides process evenly.

That blind spot is not an oversight; it follows from the doctrine's structure. When procedure is set by rule, categorical evaluation makes sense, because the rule fixes what every claimant

---

[112] Mathews v. Eldridge, 424 U.S. 319, 335 (1976). The *Mathews* inquiry evaluates the procedures attached to a program or a class of decisions, not the procedures an individual adjudicator chose to provide in a particular case.

receives. But procedural discretion unfixes it. When the governing materials tell the judge what she may do rather than what she must, the operative procedure is set not by the rule but by the judge, and the distribution of process across judges becomes a fact about the system that no categorical inquiry will capture.[113] A respondent who draws a judge at the low-process end of the distribution may receive markedly less opportunity to obtain counsel, to develop a record, and to

be heard than a respondent who draws a judge at the high-process end, and on the categorical view, both have received the process the system formally provides.

The point is not that the Constitution requires every immigration judge to grant the same number of continuances. It is that the values due process is meant to protect, accuracy, participation, the dignity of being heard before the state removes a person from the country, are unevenly delivered in a regime of procedural discretion, and that the unevenness is invisible to the doctrine that is supposed to guard those values. If the empirical findings hold, the due process question in immigration adjudication is not only whether the system provides constitutionally adequate procedures in the aggregate. It is whether a system can be said to provide due process when the process a person receives depends so heavily on an assignment she did not choose and cannot predict.[114]

[113] This is the doctrinal consequence of the procedural-discretion account developed in Shah, *supra* note 6 (manuscript at 102–05). Where procedure is discretionary rather than rule-bound, evaluating the rule tells one little about the process claimants actually receive.

[114] I do not argue that current doctrine supplies a remedy; I argue that the problem sits in the doctrine's blind spot and that the institutional reforms discussed below, rather than constitutional litigation, are the more promising response. This is consistent with the institutional, rather than constitutional, orientation of the procedural-discretion literature.

D. Consistency and the Rule of Law

The rule-of-law concern that animated *Refugee Roulette* was outcome disparity: like cases coming out differently depending on the judge.[115] Procedural variation presents a parallel concern that the outcome-focused literature has not named: like cases being *processed* differently depending on the judge. The two are related but not identical, and the procedural version is in some respects more troubling.

Outcome disparity, however objectionable, at least reflects disagreement about the thing the system exists to decide, whether relief is warranted. Reasonable adjudicators can disagree about whether a particular applicant has shown a well-founded fear of persecution, and some

---

outcome variation is the price of committing such judgments to human adjudicators at all.[116] Procedural variation is different. When two judges grant sharply different numbers of continuances to similar respondents, or resolve sharply different shares of their cases by default rather than on the merits, the divergence is not disagreement about the merits, it is divergence in the conditions under which the merits are reached. A respondent denied the continuance she needed to find a lawyer has not lost a contested legal question. She has been placed, by the accident of assignment, in a worse position to litigate whatever question her case presents.

There is a further reason procedural variation is harder to defend than outcome variation. Outcome variation is at least visible to the system that produces it: grant rates are reported, compared, and debated, and a judge whose grant rate is an outlier can be identified and asked to explain. Procedural variation has been invisible, because no one has measured it, and what is

[115] Ramji-Nogales et al., *supra* note 2, at 378–79.

[116] Some degree of outcome variation is inherent in any system that assigns contested judgments to individual decisionmakers, and not all of it is illegitimate. The procedural-variation problem is different because the variation is in the conditions of adjudication rather than in the contested judgment itself.

invisible cannot be corrected.[117] A judge whose in absentia rate is three times the rate of the judge down the hall has not, until now, been an identifiable outlier on any dimension the system tracks. Measurement is the precondition for accountability, and the procedural dimension of immigration adjudication has lacked even that.

There is also a predictability cost that compounds the consistency problem. A lawyer advising a client about whether to pursue a claim, whether to seek a continuance, or whether to risk a merits hearing must form some expectation about how the proceeding will unfold, and that expectation rests on assumptions about the process the court will provide. When procedural practice varies sharply and invisibly across judges, those assumptions become guesses, and the

quality of legal advice degrades for reasons that have nothing to do with the law and everything to do with the assignment. The respondent who cannot predict whether she will be granted the time to gather evidence cannot plan how to present her case, and the lawyer who cannot predict whether a judge will default a client who misses a hearing cannot counsel that client about the stakes of an absence. Predictability is itself a rule-of-law value, and procedural variation erodes it in a way that outcome variation, at least when grant rates are published, does not.

E. Accountability and the Direction of Political Pressure

Procedural discretion also complicates the account of political control over immigration adjudication that the prior literature developed. Kim and I showed that removal outcomes shift with presidential administration, which suggested that political pressure operates on the merits.[118] Procedural discretion offers a second, quieter channel through which such pressure can operate,

[117] The invisibility is precisely what this Article's measurement is meant to cure. A judge's continuance rate, hearing intensity, and in absentia rate can be computed and compared just as her grant rate can, and Part IV's measures supply the metrics for doing so.

[118] Kim & Semet, *supra* note 2, at 620–33.

and one that is even harder to police. An administration that wishes to increase removals need not instruct judges how to rule. It can alter the procedural environment, narrow the continuance standard, impose completion quotas, restrict administrative closure, and let the resulting compression of process do the work.[119] Process curtailed is, predictably, relief reduced and removal eased, without any visible instruction on the merits and without any decision a reviewing court is likely to disturb.

This is the troubling reading of the findings. Procedural discretion gives the political branches a means of influencing adjudicatory outcomes that is largely unreviewable, because courts review merits decisions and the reasons given for them, not the antecedent allocation of process.[120] An administration that leans on the procedural lever leaves fewer fingerprints than one that leans on the merits.

Yet there is a less troubling reading, and intellectual honesty requires putting it alongside the first. Some procedural variation reflects not political manipulation but legitimate management of a system under genuine strain. The immigration courts carry a backlog measured in the millions, and a judge who moves cases efficiently, declines unfounded continuances, and enforces appearance requirements may be serving the system's interest in timely adjudication, which is itself a component of justice for the respondents waiting years for a hearing.[121] On this reading, between-

[119] The Attorney General's authority to set binding procedural policy through certified decisions and through performance metrics gives the political leadership of DOJ a procedural lever that does not require directing outcomes. See Matter of L-A-B-R-, 27 I. & N. Dec. 405 (A.G. 2018); NAT'L ASS'N OF IMMIGR. JUDGES, *supra* note 44. The policy-period analysis in Part IV.E is designed to detect the procedural footprint of such interventions.

[120] See Shah, *supra* note 6 (manuscript at 102) (noting that procedural discretion and its downstream effects escape judicial review).

[121] See STRAUT-EPPSTEINER, *supra* note 4, at 1. Delay is not costless to respondents; prolonged pendency can mean prolonged detention, prolonged uncertainty, and the erosion of evidence. A judge who reduces delay is not obviously doing respondents a disservice.

judge procedural variation is the ordinary residue of committing case management to human adjudicators, no more sinister than the variation among trial judges in any busy court system. The difficulty is that the data cannot, by themselves, tell us which reading is correct in any given instance, and the same continuance denial can be sound docket management in one case and the curtailment of a meritorious claim in another. What the data can do, what this Article's measurement makes possible, is bring the variation into view so that the institutions responsible for the courts can ask, judge by judge and policy by policy, which reading fits. That, and not a verdict delivered from the data alone, is the contribution measurement makes to accountability.

F. Reform

The reforms that follow share a premise: procedural discretion is not an evil to be abolished but a feature to be managed, and the institutions best positioned to manage it are the ones that create it. Eliminating procedural discretion is neither possible nor desirable, because the flexibility to tailor

process to the case is part of what allows an overburdened system to function and part of what allows a conscientious judge to do justice in the hard case.[122] The goal is to constrain the unjustified variation while preserving the justified flexibility, and that is a task for institutional design rather than for prohibition.

EOIR and the Department of Justice. The agency that runs the immigration courts already collects the data this Article analyzes, and it could turn that data inward. EOIR should measure procedural style as a matter of routine, computing for each judge the continuance, hearingintensity,

[122] The argument is not for the elimination of procedural discretion but for its management, consistent with the procedural-discretion literature's own resistance to one-size-fits-all proceduralism. See Shah, *supra* note 6 (manuscript at 102) (rejecting "a one-size-fits-all approach to proceduralism" in favor of institutional tools for managing discretion).

pace, and in absentia measures developed here, adjusted for case mix and court, and using them not to discipline judges toward a single number but to identify outliers whose practice warrants review.[123] A judge whose in absentia rate sits far above her peers', after adjustment, is not necessarily doing anything wrong, but she is doing something different, and the difference is worth understanding. The measurement that makes outcome outliers visible can make procedural outliers visible too, and an agency serious about consistency would want to see both.

The agency should also recognize that its own policy levers shape procedural variation, and use them deliberately. The completion quota is the clearest example. If quotas compress process by rewarding speed over development, a question Part IV's policy-period analysis is designed to answer, then the agency has been generating procedural curtailment as a byproduct of a productivity metric, and it should weigh that cost explicitly rather than treating throughput as an unalloyed good.[124]

Measurement should also be made public, not held internally. EOIR already publishes case-completion and grant-rate statistics, and the same disclosure could extend to the procedural measures, reported at the court level and, with appropriate care, at the judge level.[125] Public

[123] The agency is better positioned than any external researcher to construct these measures, because it holds the complete and current data and can adjust for case characteristics it does not release. The recommendation is that it construct and use the measures internally, with the adjustment for case mix and court that distinguishes genuine outliers from judges with unusual dockets.

[124] The recommendation is not necessarily to abolish completion goals but to set them with awareness of their procedural footprint, informed by the kind of before-and-after analysis Part IV.E describes. The policy-period analysis is consistent with quota-driven compression of process: the between-judge standard deviation of pace fell 14% after the quotas took effect, even as the in absentia share rose.

[125] EOIR's existing public reporting of completion and outcome statistics provides a precedent for publishing procedural measures. Disclosure would let respondents, their representatives, and the public see the procedural

reporting would do two things the existing statistics do not. It would let a respondent and her representative know, before a hearing, something about the procedural environment they are entering, much as published grant rates already inform expectations about outcomes. And it would expose the measures to outside scrutiny, allowing researchers and advocates to test the agency's construction and to identify patterns the agency might prefer not to see. Transparency about process is a modest reform, and it does not by itself change a single procedural choice. But it is the precondition for every other reform, because a variation that no one outside the agency can observe is a variation that no one outside the agency can challenge.

The Board of Immigration Appeals. The Board reviews immigration judge decisions, and it is the natural locus for developing a more articulated law of procedural discretion. At present, the "good cause" standard for continuances and the standards governing in absentia orders are thin, and their content is supplied case by case, which is part of why practice varies so widely.[126] The Board could narrow the range of defensible practice, not by mandating a fixed number of

continuances, but by articulating the considerations that bear on the procedural choices judges make, so that a judge's exercise of procedural discretion becomes reviewable against a standard rather than essentially unreviewable. A reviewable standard would do for procedure what merits review already does for outcomes: subject the judge's choice to a check.

The federal courts. Article III courts have largely treated procedural choices as committed to the immigration judge's discretion and unreviewable, and the structural account in this Article

dimension of the courts that the agency's internal review would address, and it would let external researchers test and refine the measures. See EXEC. OFF. FOR IMMIGR. REV., *supra* note 111.

[126] The continuance standard, 8 C.F.R. § 1003.29, supplies no definition of "good cause," and the resulting gap is filled by individual judges and by episodic Attorney General intervention. A more developed body of Board precedent would narrow the range of defensible practice without eliminating discretion.

suggests why that posture is too deferential.[127] When a continuance denial or an in absentia order is the proximate cause of a removal reached without counsel, without evidence, and without a merits hearing, the procedural choice has done the work the merits decision is normally thought to do, and it deserves the scrutiny a merits decision would receive. Reviewing courts need not secondguess every scheduling order. They could, however, treat patterns, a record showing that a removal followed a denial of the time a respondent needed to obtain the counsel that the data show changes outcomes, as raising a question the law should be willing to ask.

Congress. The deepest source of procedural variation is statutory: the immigration laws commit vast procedural authority to adjudicators while supplying almost no procedural floor, and they place those adjudicators inside an enforcement-oriented department.[128] Congress could supply a procedural floor, minimum continuance entitlements for respondents seeking counsel, limits on in absentia orders where notice is contested, a guaranteed opportunity to develop a record in cases seeking relief, that would convert some of today's discretionary practice into entitlement, compressing the variation at its lower end. The longstanding proposal to remove the immigration

courts from DOJ and reconstitute them as an independent Article I court speaks directly to the accountability concern: an adjudicator insulated from the enforcement priorities of the political branches is less available as an instrument of the quiet, procedural form of political control identified above. Whether to take that step is a contested question of institutional design on which reasonable people disagree, and I do not claim the data resolve it. But the procedural findings

[127] The courts' general reluctance to review procedural choices in immigration adjudication leaves procedural discretion as the least-checked form of immigration-judge authority. See Shah, *supra* note 6 (manuscript at 102, 145).

[128] The structural placement of immigration judges within DOJ, and the absence of a statutory procedural floor for removal proceedings, are choices Congress made and could revisit. Proposals for an Article I immigration court have circulated for years and bear directly on the procedural-discretion problem. See, e.g., Kim, *supra* note 19, at 50–55 (discussing structural reform options).

supply a consideration the debate has lacked: the structural placement of immigration judges shapes not only how they decide but how much process they provide, and the cost of the current structure is paid partly in process, unevenly distributed, and largely unseen.

G. Beyond Immigration

The immigration courts are this Article's subject, but they are not the only place procedural discretion operates, and the findings carry a lesson for administrative adjudication generally. Most federal adjudication is informal, conducted outside the APA's formal-hearing model by adjudicators who, like immigration judges, manage process under conditions of high volume and thin procedural constraint.[129] Social Security disability hearings, veterans' benefits adjudications, and the benefits determinations of other mass-justice agencies share the structural features that make procedural discretion possible in immigration court: individual adjudicators, enormous dockets, recorded events, and procedures that the governing materials commit to discretion rather than fix by rule.[130] In each, the disposition-focused research tradition that has dominated immigration scholarship is likely to have overlooked the same procedural dimension this Article

---

recovers, and the same methodology, measuring adjudicator-level procedural style from administrative event records, could recover it.

The broader claim is methodological as much as substantive. Administrative law has theorized procedural discretion but has lacked a way to see it, because procedure does not announce

[129] See *supra* Part I.B; Shah, *supra* note 6 (manuscript at 101–02). The procedural-discretion account is explicitly cross-agency, encompassing the National Labor Relations Board, the State Department, and the consumer-protection agencies, among others.

[130] On the disability adjudication system as a comparable mass-justice setting with documented adjudicator variation, see JERRY L. MASHAW, BUREAUCRATIC JUSTICE (1983); see also David Ames et al., *Due Process and Mass Adjudication: Crisis and Reform*, 72 STAN. L. REV. 1 (2020) (documenting variation across Social Security administrative law judges).

itself the way an outcome does. A grant or a denial is a discrete, recorded event with a known valence; a continuance practice or a default rate is a pattern that becomes visible only when the events are aggregated to the adjudicator and compared. The contribution of an empirical study like this one is to make the pattern visible, and the contribution generalizes: wherever an agency adjudicates at scale and records what it does, the procedural style of its adjudicators can be measured, and the uneven distribution of process that procedural discretion produces can be moved from argument onto an evidentiary footing. Immigration court is where I have done this. It is not the only place it can be done.

Conclusion

For two decades, the study of immigration adjudication has been a study of disposition. We have asked who grants and who denies, and we have learned, to a degree that should trouble anyone committed to the rule of law, that the answer turns on the judge a respondent happens to draw. This Article has asked a prior question. It has asked not how immigration judges decide but how they manage the process that precedes the decision, whether they differ in the counsel they enable, the access to a hearing they grant, the speed and attention they provide, and the procedural exits on which they rely, and it has shown that this procedural dimension of judging is real, measurable, persistent, and consequential.

The result that should change how the field reads its own findings is the last one. Judges differ in continuance practice, hearing intensity, and pace; that much most observers would have expected. The unexpected finding is that a substantial share of what scholarship has attributed to adjudicator disagreement about the merits, on the order of a quarter to a third of the between-judge variation in case outcomes, is associated instead with how judges manage process. Some of the refugee roulette is a procedural roulette. The disparities that two decades of research have traced

to ideology, background, and legal disagreement are, in meaningful part, disparities in the path a case takes to its decision rather than in the decision itself. That is not a refinement of the disparity literature; it is a partial reinterpretation of it, and it points future work toward a determinant of immigration-court outcomes that has been recorded in the data the whole time and never assembled into view.

The reinterpretation cuts in two directions. It is troubling, because procedural discretion gives the conditions of justice a randomness the system does not acknowledge and that current doctrine cannot see, and because it offers the political branches a quiet, largely unreviewable channel for shaping outcomes by shaping process. It is also, on a fairer reading, the ordinary residue of asking human adjudicators to manage an impossible volume of cases, some more generously than others for reasons that have nothing to do with manipulation and everything to do with judgment under strain. The data do not choose between these readings. What measurement offers is the capacity to ask the question case by case and judge by judge, to identify the procedural outlier as we already identify the substantive one, and to subject the allocation of process to the scrutiny we have long reserved for the allocation of outcomes.

The immigration courts provide an unusually clear window onto a phenomenon that administrative law has theorized but never observed at scale. What is visible through that window is a system in which the process a person receives before being removed from the country, and with it some part of her chance to remain, depends on an assignment she never made. Whatever one concludes about the right balance between flexibility and consistency in the administration of justice, that fact belongs in the debate, and the burden of this Article has been to put it there with evidence rather than assertion.