

[v 2.0 (Draft; please don't quote without permission from the author, his former roommate's two cousins, and at least one dog)]

WHY JUDICIAL REVIEW FAILS: ORGANIZATIONS, POLITICS, AND THE PROBLEM OF AUDITING EXECUTIVE DISCRETION

Mariano-Florentino Cuéllar *

Every day executive branch officials make thousands of decisions affecting our security and welfare. Homeland security officials screen tens of thousands of people at the border. They decide whose name gets on government “no fly lists.” Agencies freeze suspected terrorist assets, choose what companies to inspect for environmental violations, and decide whom to prosecute. This article describes how judicial review predictably and systematically fails to prevent abuse and promote organizational learning when government officials make many such choices using their discretion to target individuals or groups. It then proposes the use of quasi-judicial audits of executive discretion as a remedy. While it is rare that discretionary decisions are entirely immune from some kind of judicial review, courts’ role is often so circumscribed or deferential that the probability of uncovering problems almost certainly falls close to zero. The resulting amount of executive discretion carries considerable risks along with rewards. Some decisions no doubt benefit from the speed and accountability that results from limiting judicial intervention. Yet judicial review’s evisceration probably makes it easier for some government officials to subtly manipulate their discretion to promote appealing political impressions, for others to engage in outright malfeasance, and for still other (more virtuous) officials to simply fail to learn from their mistakes – whether these arise in deciding who to charge with a federal crime, who to designate as an enemy combatant, or how much money to freeze in a suspicious charity’s account. The reliance on judicial review to manage discretion makes it hard to address these concerns in part because courts routinely define much of their work in terms of applying the same standard of deference to every case in a particular class, making it difficult to increase the stringency of review in some policy domains without making the costs allegedly prohibitive. As a conceptual alternative, I propose a framework for systematically auditing samples of discretionary decisions and making those results public. Audits help sever the connection between the perceived costs of encroaching on discretion and the stringency of review, and avoid the potentially distorted picture of bureaucratic activity created by a litigation-driven process. These properties make audits a promising supplement to judicial review in those instances where it is plausible to believe that more could be learned from incisively studying a subset of cases instead of superficially reviewing more of them. Despite their potential value, such audits are almost never done by existing federal audit bureaucracies (the congressional Government Accountability Office and the department-specific Inspector General Offices), nor does the legislature seem to do them itself in connection with oversight hearings. I conclude by discussing some of the political and bureaucratic dynamics working against these audits and suggesting how they may be weakened.

INTRODUCTION	2
I. AUDITS ENHANCE THE SUPERVISION OF EXECUTIVE DISCRETION	7
a. <i>Executive Discretion is Pervasive</i>	8
b. <i>Discretion is Routinely Managed Through Variations in Standards of Judicial Deference</i>	10
c. <i>Audits Can Substitute For, or Supplement, Judicial Review</i>	11
d. <i>Traditional Judicial Review Suffers From Limitations</i>	17
e. <i>Examples</i>	21
i. <i>Freezing Assets of Suspected Terrorists</i>	21
ii. <i>Government-Run “No-Fly” Lists</i>	24
iii. <i>Environmental and Occupational Safety Inspections</i>	25

* Associate Professor and Deane F. Johnson Faculty Scholar, Stanford Law School; Faculty Affiliate, Stanford Center for International Security and Cooperation. I appreciate helpful conversations with Glenn A. Fine and Daniel Sutherland, as well as with Mark Kelman, Pam Karlan, Barbara Fried, Deborah Hensler, Rachel Barkow, Mark Lemley and Jeff Strnad. I am especially indebted to Richard Craswell and Alan Morrison.

	f. <i>Audits Can Ameliorate the Limitations of Traditional Judicial Review</i> ..	27
II.	THE INSTITUTIONAL DESIGN PROBLEMS ARE TRACTABLE	29
	a. <i>Audits Can Be Adapted to Address Multiple Forms of Discretion</i>	29
	b. <i>Most Institutional Design Problems Have Plausible Solutions</i>	30
III.	REFORMS CONTRAST SHARPLY WITH THE STATUS QUO	34
	a. <i>Federal Bureaucracies Do Not Ordinarily Perform Audits of Executive Discretion</i>	34
	b. <i>Neither Do Legislators</i>	36
IV.	AUDITS FACE POLITICAL CONSTRAINTS	38
	a. <i>Observers Sometimes Fail to Distinguish Between Targeted Discretion and Policy Judgments</i>	38
	b. <i>Political Actors Have Polarizing Incentives</i>	39
	c. <i>Institutional Inertia Locks In Existing Conceptions of Adjudication</i>	41
	d. <i>Narrow Conceptions of Auditor Mission Persist</i>	41
CONCLUSION	43

INTRODUCTION

The problem is a familiar one. Government officials have staggering discretion when deciding what industrial plants to inspect for safety violations, whom to indict, who gets turned away by the Department of Homeland Security (DHS) at a port of entry without being allowed to plead her case for asylum, what allegedly terrorist assets to freeze, or whose name DHS puts on a government “no fly” list. Despite an often-mentioned social commitment to judicial restraints on public power,¹ our laws routinely create and protect discretion instead of restraining it. Which means public officials making certain crucial discretionary choices are, in turn, rarely restrained by courts — or by anything except their own conscience or occasional fits of political attention. All this discretion presumably allows executive authorities to make decisions quickly and efficiently, and the public to hold those authorities accountable. It spares executive authorities the worry that courts will endlessly second-guess their decision when they decide what target to bomb or what assets to freeze. Unfortunately, discretion also lets authorities shirk instead of learning from their mistakes, and abuse their powers instead of using them responsibly. Which raises the pressing question of how public organizations with such considerable discretion will (be forced to) learn from their

¹ See, e.g., *Citizens to Preserve Overton Park v. Volpe*, 401 U.S. 402 (1971) (subjecting an informal, discretionary decision of the Secretary of Transportation to judicial review on the basis of statutory language prohibiting federal aid for highways through public parks unless “no feasible and prudent alternative” existed). *Overton Park* set the stage for a substantial expansion in the availability (and stringency) of judicial review governing informal, discretionary decisions. But review remains either unavailable or fairly cursory for a massive range of discretionary decisions involving national security, foreign policy, immigration, domestic regulatory enforcement, public benefits, and investigation or prosecution. See KENNETH CULP DAVIS, *DISCRETIONARY JUSTICE* (1969). Regarding the trope that judicial review should have an exalted role in constraining arbitrary bureaucratic action, see Rachel E. Barkow, *More Supreme Than Court? The Fall of the Political Question Doctrine and the Rise of Judicial Supremacy*, 102 COLUM. L. REV. 237 (2002); Lisa Schultz Bressman, *Beyond Accountability: Arbitrariness and Legitimacy in the Administrative State*, 78 N.Y.U. L. REV. 461 (2003); David Cole, *Judging the Next Emergency: Judicial Review and Individual Rights in Times of Crisis*, 101 MICH. L. REV. 2565 (2003); James E. Pfander, *article I Tribunals, article III Courts, and the Judicial Power of the United States*, 118 HARV. L. REV. 643 (2004).

mistakes, and how they will be policed against abusing their legal powers. Although judicial review is often considered a central tool in preventing mistakes or abuses,² the benefits of discretion have led to a vigorous doctrinal and policy debate about the proper stringency of such review.³

This article challenges the terms of that familiar debate in order to address the underlying question of how discretion should be supervised in a complicated political environment. It tells the story of how judicial review systematically fails to effectively balance discretion's costs and benefits, particularly when it comes to certain discretionary choices like the ones in the first paragraph, where the executive branch has special powers to significantly affect the fate of an individual, group or company.⁴ In the process, I show how the paradigm of judicial review often ill-serves the valuable goals of helping bureaucratic organizations learn from their failures and avoid political pressures that endanger their missions. The problems arise both in national security and domestic regulatory contexts, domains that have been traditionally treated separately but increasingly blur.⁵ The article then shows how *audits* of targeted discretion can fill certain gaps left by deferential judicial review, or indeed, by the absence of any judicial review at all.

Part I begins by reviewing the extent of executive discretion, and the traditional approach courts and legislators use to control it. What this brief examination highlights is how no plausible justification for executive discretion can deny the existence of a trade-off between the value of discretion and its potential cost. Bureaucratic organizations need some flexibility to do their job. On the other hand, any meaningful review process consumes resources, takes time, and reduces the freedom of choice that makes executive discretion valuable in the first place. In response to this dilemma, courts routinely apply differing degrees of stringency when reviewing executive decisions. Their pervasive calculus of deference may take shape at legislators' explicit behest (as with the CIA Director's power to decide which employees are national security risks) or at the courts' own hands. Consider, for example, the fate of individuals that executive

² See, e.g., *City of Chicago v. Int'l College of Surgeons*, 522 U.S. 156, 184 (summarizing previous interpretation of federal jurisdictional statute by emphasizing the Court's conclusion that "Congress meant to hold federal agencies accountable by making their actions subject to judicial review."); Rebecca L. Brown, *Accountability, Liberty, and the Constitution*, 98 COLUM. L. REV. 531, 535 (1998)(emphasizing the alleged role of judicial review in promoting accountability).

³ Voices on one side emphatically insist on greater opportunities for highly-stringent judicial review of executive branch actions. See, e.g., Culp Davis, *supra* note 1, at 216 ("The vast quantities of unnecessary discretionary power that have grown up in our system should be cut back, and the discretionary power that is found to be necessary should be properly confined, structured, and checked"). Cole, *supra* note 1, at 2567. Similarly emphatic voices take the position in equipoise. See, e.g., Eric A. Posner and Adrian Vermeule, *Accommodating Emergencies*, 56 STAN. L. REV. 605 (2003); Ruth R. Wedgwood, *Al Qaeda, Terrorism, and Military Commissions*, 96 AM. J. INT'L L. 328 (2002). Similar debates play out in the context of constitutional torts. See, e.g., James J. Park, *The Constitutional Tort Action As Individual Remedy*, 38 HARV. C.R.-C.L. L. REV. 393, 395 (2003). At least some of the debate turns on differing views about the extent to which a larger "political process" promotes "accountability." I discuss this in Parts III, and IV.

⁴ See Part I.a. *infra*.

⁵ Cf. Cass R. Sunstein, *Administrative Law Goes to War*, 118 HARV. L. REV. 2663, 2672 (2005)("In war no less than in peace, the inquiry into presidential authority can be organized and disciplined if it is undertaken with close reference to standard principles of administrative law.").

authorities designate as enemy combatants. Courts can increase the stringency of review by requiring more thorough hearings before someone is designated, and by decreasing the deference accorded to the outcome of those hearings or (in the absence of hearings) to the executive determinations themselves.⁶ Greater stringency of review presumably reduces the probability that someone would be labeled an enemy combatant by mistake, or because she is (for example) merely a political opponent of the president. But it allegedly increases the resources that society must expend on the review process and that the executive branch must expend defending its decision.⁷ If stricter review consumes substantially greater resources or creates a material possibility of embarrassment for executive officials, it may also chill the authorities from designating individuals that should (in an ideal world) receive such a designation. The same problem arises in the other contexts involving targeted discretion, such as freezing assets and making prosecutorial decisions. In response, courts tend to vary the deference they give the executive or her agent.

Yet this account neglects an alternative way of managing the costs and benefits of discretion besides adjusting the stringency or availability of judicial review. The alternative is for some court-supervised or independent authority to use *audits* to sample discretionary decisions and to review them more thoroughly, rejecting the assumption that a given degree of review stringency should be applied to all cases in a class. Despite their relative absence from discussions of how to constrain government discretion, audits of this kind are familiar from a panoply of private and public sector contexts – as with insurance companies that review a small slice of their closed files to learn from their past behavior, tax enforcers who audit a sample of the public at random (or nearly so), or courts handling complicated class actions and picking out a sample of cases from the class to learn more about the merits. Because they are an alternative to imposing a single

⁶ Compare *Webster v. Doe*, 486 U.S. 592, 602-04 (1988) (finding CIA director's power to fire employee on national security grounds committed by law to agency discretion) with *Klamath-Siskiyou Wildlands Center v. Bureau of Land Mgt.*, 387 F.3d 989 (9th Cir. 2004) (finding, under an arbitrary and capricious standard that the court understood to require "hard look" review, that the Bureau of Land Management's environmental assessments of two timber sales, conducted pursuant to the National Environmental Policy Act, were inadequate because they failed to consider the cumulative impact of the sales). I don't mean to minimize the subtleties of the variegated constitutional, statutory, and prudential doctrines on which courts (and even legislatures) draw when they decide on how much discretion to grant. Separation of powers, deference to national security and foreign policy decisions, judicial deference to expert determinations of government agencies, and statutory interpretation techniques all figure in this process. Even the two cases I cite here represent extraordinarily different contexts, and the kinds of discretion involved in the decision are also different. The point here is how nearly any plausible applications of such doctrines require (or at least allow) some consequentialist balancing of the costs and benefits associated with discretion, and different ways of striking that balance are associated with distinct degrees of stringency in the court's review of some executive decision. As these two cases show, courts indeed strike different balances when applying these doctrines, and in the process, they set different degrees of stringency for the review of discretionary executive decisions.

⁷ See, e.g., Reply Brief for the Petitioner, Supreme Court of the United States, *Heckler v. Chaney*, 1984 WL 566059, 4 (November 23, 1984) ("[R]espondents' submission, if accepted, would allow *anyone* to seek judicial review of the agency's decision not to bring enforcement proceedings under any portion of the Act.") (emphasis added); Brief for the Respondents, Supreme Court of the United States, *Hamdi v. Rumsfeld*, 2004 WL 724020, 12 (March 29, 2004) (arguing that further factual development of the circumstances surrounding an alleged enemy combatant's designation as such "would divert the military's attention from the ongoing conflict in Afghanistan...").

standard across the board, audits have the potential to disrupt the familiar, repetitive debate about whether our society deserves greater court protection of its rights and prerogatives. Put differently, *even* if one accepts the executive branch's strident (and frequently questionable) assertions that the sky would fall if discretion were more easily reviewed in court, there remains a viable option for reviewing that discretion without incurring the various costs associated with traditional judicial review.

This wouldn't matter much if the conceptual choice between deferential review and audits were a mere quirk of form. Not so. Under quite reasonable assumptions, deferential judicial review could be worse than review using audits even though the costs may be similar. The absence of audits, conversely, diminishes our system's capacity to detect executive branch manipulation.⁸ It may also dampen the incentives of executive branch bureaucracies to learn from their mistakes, and makes it easy for key actors in the system to avoid articulating (either in statutory or executive mandates) what standards are actually supposed to govern executive discretion. Together these dynamics ultimately affect the costs and benefits of laws that grant the executive branch discretion in the first place, and also the political context governing those grants of legal power. In principle, we should expect an accountability-power trade-off: some people and groups should be willing to see the executive branch get more power but only if it could be reliably supervised. While the devil may be in the details, Part II surveys some of these problems and discusses how they might be plausibly resolved.

Little of this thinking has influenced the review of the executive branch's discretionary decisions. Instead, executive discretion is supposed to be policed through two standard mechanisms quite unlike the audits I described above. At one end of the spectrum lie courts that seem to conceive of the task at hand in terms of articulating or (on the basis of legislative enactments) applying different standards of deference to review all discrete decisions of a particular kind (i.e., all enemy combatant designations, all personnel decisions in intelligence agencies, all decisions by OSHA to initiate investigations, and so on). At the other end lie mechanisms that combine some kind of external inspection of agency practices with political safeguards: Inspector General's Offices, the Government Accountability Office, congressional oversight, and similar procedures that produce reports on financial management, discussions of organizational problems, hearings, and testimony. While the court-centered end of the spectrum suffers from the aforementioned limitations, one might imagine that review provided by audit bureaucracies and the legislatures could fill the breach. As Part III shows, it's not that simple. Rarely if ever do these mechanisms focus on substantively evaluating a discrete sample of targeted executive decisions rather than the finances or performance of entire programs, and some kinds of decisions (like prosecutorial charging decisions) seem a

⁸ Of course, the mere creation of some auditing system does not automatically solve organizational learning and accountability problems involving the law. As I note in Part II, a great deal depends on details of institutional design. The impact of an audit system also depends on the public's response, and the institutional dynamics affecting that response. Audit systems can be counterproductive if they merely provide a false sense of security – which is in some sense precisely my criticism of judicial review in many of the contexts I discuss in this paper. Nonetheless, the status quo seems even more likely to provide precisely that false sense of security because it lacks many of the potential advantages that a carefully-structured audit system could generate. For a thoughtful discussion of role of audits and the pitfalls in designing them, see MICHAEL POWER, *THE AUDIT SOCIETY: RITUALS OF VERIFICATION* (1999).

trifle too mundane to attract much attention from our existing auditing bureaucracies. In fact, using new data compiled from samples of reports from the GAO and Inspectors General, I show how these audit bureaucracies do nearly anything *except* the audits of executive discretion I advocate here.

Why, then, have audits of targeted discretion remained largely irrelevant in reviewing executive discretion? In Part IV I discuss four reasons, which together help make a more general point about the likely disconnection between what we expect of our public law and what public bureaucracies actually deliver. The first is a frequent failure to distinguish between different types of executive discretion that may call for distinct kinds of review. “Targeted” discretionary decisions (applying an implicit or explicit legal standard) are most obviously suited for audit review. Broad policy judgments that lie at the core of regulatory rulemaking, on the other hand, may be harder to review in this fashion. Second, key players may lack an organizational or policy interest in audits because they have an investment in perpetuating some aspect of the current system. The deference approach may yield greater discretion on balance to the executive because it makes reviewing authorities deferential across the board. On the other side, advocates of restricting discretion may also go for broke through legal and legislative strategies that apply across the board, and may galvanize supporters with troubling scenarios. Third, the current system is boosted by some inertia borne from adversarial adjudication, where lawyers tend to conceive of their role as zealous advocates on behalf of individual clients and reviewing institutions (especially courts) are conceived as direct protections of the rights of similarly-situated individuals. Finally, the existing audit bureaucracies fail to use their legal authority to do audits of targeted executive discretion in part because of a persistent preoccupation with the sort of financial auditing mission that initially led to their creation.

Together these factors may lock in suboptimal institutional responses to serious legal problems. Through this case study and thought experiment I hope to highlight three crucial challenges that follow from that sort of lock-in, and that may be relevant to a large class of problems: (1) the importance of recognizing the inherent limitations of judicial review as a means of managing government discretion, (2) the value of envisioning new institutional designs to manage discretion more effectively, and (3) the need for reasonable strategies to implement those designs in a politically complicated world. No one should underestimate judicial review’s enduring value. Instead, it is the stale discussion of its merits as a means of promoting accountability that must be transcended. Perhaps the most striking thing about executive discretion is that the problem of trading off its costs and benefits is so pervasive, yet the means of balancing those costs and benefits has remained unduly narrow. This article begins the task of broadening the scope of potential solutions and understanding the forces that shape our perceptions of whether those solutions even exist.

I.

AUDITS ENHANCE THE SUPERVISION OF EXECUTIVE DISCRETION

Ordinary citizens tend to carry around a few simple mental narratives about the stakes involved when government makes such discretionary decisions.⁹ One of those narratives recurs perhaps more often than any, and goes something like this. Government bureaucracies get power from legislators, the “people” who elect the legislators, or maybe even (among the sporadic, more darkly cynical observers) a shadowy conspiracy of oligarchs. The frontline employees and officials who work in government bureaucracies then use their power to regulate working conditions, protect the air, screen containers, safeguard the borders, or to accomplish a host of other goals that the principals care about and could not be achieved without giving the government such powers.¹⁰ But since that power can be abused, those to whom it is entrusted must be policed to make sure they adhere to the proper (legal) rules and standards that are supposed to govern their decisions.¹¹ In societies like our own, courts are assumed to play a particularly important role in doing that policing.¹² Our public credo makes much of how courts promote accountability. They scrutinize instances where government uses discretion to target an individual or group. In doing so, they are assumed to help the public learn what ails the government’s efforts to protect our welfare and security.¹³ They’re routinely assigned responsibility for ensuring that people’s rights are respected, that agencies don’t exceed their legal powers when they promote security or safety, and that the law is correctly applied to the millions of people and organizations who are affected by government every day.¹⁴

The preceding assumptions about the role of courts are, of course, wrong. The law may commit a decision to agency discretion. Courts may be forced by law or custom to review executive branch actions only with extreme deference. Judges may lack jurisdiction. In short, when courts weigh in on a great many discretionary decisions to make sure they’ve been done right, they may have little choice but say: “Glad you’re doing such a great job.”¹⁵ Sometimes legislators explicitly confer enormous discretion,

⁹ See Mark Schlesinger and Richard R. Law, *The Meaning and Measure of Policy Metaphors*, 94 AM. POL. SCI. REV. 611 (2000).

¹⁰ See, e.g., B. Dan Wood, *Principals, Bureaucrats, and Responsiveness in Clean Air Enforcements*, 82 AM. POL. SCI. REV. 213 (1988); Geoffrey Brennan and Alan Hamlin, *On Political Representation*, 29 BRITISH J. POLI. SCI. 109 (1999).

¹¹ See Barry R. Weingast, *Political Foundations of Democracy and the Rule of Law*, 91 AM. POL. SCI. REV. 24 (1997).

¹² See, e.g., Martin Shapiro, *Administrative Discretion: The Next Stage*, 92 YALE L.J., 1487 (1983).

¹³ See, e.g., Neal K. Katyal and Laurence H. Tribe, *Waging War, Deciding Guilt: Trying the Military Tribunals*, 111 YALE L.J. 1259 (2002); David Cole, *Judging the Next Emergency: Judicial Review and Individual Rights in Times of Crisis*, 101 MICH. L. REV. 2565 (2003); Eric A. Posner and Adrian Vermeule, *Accommodating Emergencies*, 56 STAN. L. REV. 605 (2003).

¹⁴ See generally Peter M. Shane, *Interbranch Accountability in State Government and the Constitutional Requirement of Judicial Independence*, 61 LAW & CONTEMP. PROBS. 21 (1998).

¹⁵ Courts can restrain egregious government conduct of some kinds, such as those that might give rise to viable constitutional tort claims. See *Bivens v. Six Unknown Named Agents of the Fed. Bureau of Narcotics*, 403 U.S. 388, 389 (1971). But courts proceed with extreme caution in this realm given their concern that the scope of a claim they recognize (and the remedies they might make available) would

prohibiting all or most judicial review of certain decisions, by embedding a limiting provision in an agency's legal mandate. The law says that the CIA Director, for instance, has authority to fire employees for being national security risks. It also says he has the power to define what "national security risk" means, which lets him arbitrarily fire someone for being gay (he has).¹⁶ In other cases, executive agencies have discretionary powers because prevailing statutory interpretations and constitutional provisions imply the existence of such power. That's why prosecutors have the power to choose whom to charge with little concern that their choices will be questioned in court or in any other legal forum.¹⁷ In still other cases agencies have discretionary power because courts and other external observers only review certain kinds of executive decisions with great deference, which leaves the president, the agencies he supervises, and similar executive authorities with residual control over governance.¹⁸ The resulting, legally sanctioned, executive discretion lets agencies act swiftly, learn to deal with unfamiliar situations, and (it is often believed), lets the principals hold the executive branch of government accountable for what it's done.¹⁹ But those powers carry risks as well as rewards – and striking a reasonable balance between flexibility and constraint turns out to be devilishly difficult.

A. *Executive Discretion is Pervasive*

Discretion, as I define it, is the (explicit or implicit) legal flexibility to use government power – including, but not limited to, personnel, budgets, information, and legally-sanctioned coercive authority) to affect the world. Government officials use some discretion virtually every time they do something. Though government actions are rarely purely discretionary, neither is discretion ever entirely absent. What discretion is present may come from explicit statutory language, an implicit interpretation of a statute or constitutional provision, or it could be the result of a certain institutional structure involving extremely deferential external review, which gives government officials a lot of power.

Since there's no such thing as a total absence of discretion when government pays out welfare support, prosecutes, and otherwise applies the law, the distinctions that lawyers and policymakers fight over tend to be about whether to give the executive branch relatively *more*, or relatively *less* discretion compared to a certain baseline. When the President's lawyers talk about the benefits of discretion, they're implicitly accepting a baseline state of the world where courts play a significant role in reviewing government

interfere with some of the more valuable properties associated with executive discretion. *See generally* PETER SCHUCK, *SUING GOVERNMENT: CITIZEN REMEDIES FOR OFFICIAL WRONGS* (1983).

¹⁶ *Webster*, 486 U.S. at 602-04.

¹⁷ *United States v. Armstrong*, 517 U.S. 456 (1996). As Culp Davis cogently observed, "[a] judicial trial is an acceptance of a prosecutor's decision to prosecute, not a review of it. Even a quick finding of not guilty may leave untouched the harms that flow from the prosecution." Culp Davis, *supra* note 1, at 209 n.21.

¹⁸ *See Hamdi v. Rumsfeld*, 542 U.S. 507 (2004).

¹⁹ *See* Steven G. Calabresi and Saikrishna B. Prakash, *The President's Power to Execute the Laws*, 104 *YALE L.J.* 541 (1994); Saikrishna B. Prakash and Michael D. Ramsey, *The Executive Power Over Foreign Affairs*, 111 *YALE L.J.* 231 (2001); John C. Yoo, *War, Responsibility, and the Age of Terrorism*, 57 *STAN. L. REV.* 793 (2004).

action.²⁰ That's a familiar position in the United States and in most other developed nations (and many developing ones).²¹ In criminal prosecutions, voting rights cases, and labor law injunctions, for example, some action of the executive branch (such as subjecting someone to the detriments associated with being convicted of a crime) depends on convincing a court to do something. Observers and policymakers may have different political views about how easy it should be to impose a labor injunction (for example) or convict someone of a crime. But if they don't succeed in persuading the legislature to water down the substantive standard that applies, they'll have to mount a vigorous case before a court that is quite persistently unwilling to simply defer to executive discretion. Even when such review is not occurring, the executive branch presumably labors in the shadow of the embarrassing possibility that a license grant, a regulatory rule, a criminal conviction, or a statutory enactment will be invalidated.²² This implies that we can measure the *benefits* of executive discretion against a baseline of relatively intrusive judicial review.

²⁰ See, e.g., Brief for Respondent, Supreme Court of the United States, *Hamdi v. Rumsfeld*, 2004 WL 724020 (March 29, 2004). The government's language in the brief is typical of the positions that lawyers for the executive branch have taken in this Administration – and not dramatically different (on the core issue of deference – from that taken by lawyers for other presidential administrations. It states:

As this Court has observed, courts traditionally have been reluctant to intrude upon the authority of the Executive in military and national security affairs. The customary deference that courts afford the Executive in matters of military affairs is especially warranted in this context. A commander's wartime determination that an individual is an enemy combatant is a quintessentially military judgment, representing a core exercise of the Commander-in-Chief authority. Especially in the course of hostilities, the military through its operations and intelligence-gathering has an unmatched vantage point from which to learn about the enemy and make judgments as to whether those seized during a conflict are friend or foe.

Id. at 25-26 (citations omitted).

²¹ This statement should not obscure the massive extent of variation among legal systems, many of which assign quite different roles to judicial institutions. The point is that it's quite common for those different systems to assign considerable importance to the goal of reviewing executive action through courts. See generally Anne-Marie Slaughter, *A Typology of Transjudicial Communication*, 29 U. RICH. L. REV. 99, 101 (1994).

²² Actually measuring the precise impact of review with some analytical clarity is enormously complex, but a number of scholars have made convincing arguments to this effect using qualitative or quantitative methodologies in different contexts. For some cogent examples, see JERRY L. MASHAW & DAVID HARFST, *THE STRUGGLE FOR AUTO SAFETY* (1990)(suggesting that NHTSA's reliance on costly recalls of questionable safety effects rather than prospective rulemaking has in part been driven by the impact of intrusive judicial review in rulemaking); Thomas O. McGarity, *The Role of Government Attorneys in Regulatory Agency Rulemaking*, 61 LAW & CONTEMP. PROBS. 19 (1998)(discussing the impact of the "ossification" of rulemaking, where judicial review among other factors shapes agencies' willingness to use regulatory authority); Brandice Canes-Wrone, *Bureaucratic Decisions and the Composition of the Lower Courts*, 47 AM. J. POLI. SCI. 205 (2003)(analyzing whether changes in the ideological composition of lower courts affected decisions of the U.S. Army Corps of Engineers to grant permits for development of wetlands, and finding that a standard deviation increase in estimated pro-environmental ideology of the lower courts decreased the probability that the Corps would grant a permit by 14%).

B. Discretion Is Routinely Managed Through Variations in Standards of Judicial Deference

Courts and legislatures provide for such review in accordance with a straightforward pattern. Whether because of their own decisions or because of legislative mandates, it should hardly seem surprising that courts and legislators seem to manage the costs and benefits of discretion by varying the deference they give the executive or her agent. Some decisions get more deference, and others get less.²³ As I use the concept here, deference is not quite the same thing as the “standard of review” that a court of appeals uses to review lower courts or administrative decisions, nor is it confined to courts’ review of decisions that allegedly turn on agency statutory interpretations (though it can certainly encompass this). Instead the concept applies to several considerations that together determine the stringency of a court’s review of an executive decision. That bundle of considerations includes the standards of review, the extent of required procedures a court can monitor (i.e., the “meaningful opportunity to be heard” mentioned in *Hamdi*),²⁴ and routine practices that courts use to decide on the extent of the executive branch’s flexibility to make some kinds of decisions.²⁵ Stringency of review, as I define it, encompasses among other things the standard of review governing appeals of specific administrative actions. It is affected by the degree of outright deference given to the executive branch, and the extent to which courts find through constitutional or statutory interpretation that a particular decision to be committed by law to agency discretion. More stringency means a combination of: more rigorous procedures (such as those that might be imposed on due process groups) that the government must follow before imposing a cost on someone, a less permissive standard of review for the factual findings of executive branch agencies (or lower courts), and less overall deference to the government’s decision itself. In my terms, when a court says that a six-page declaration from a Defense Department bureaucrat is enough reason to detain someone for an indefinite period of time, it is being more deferential.²⁶ When a court says that’s not enough, because the executive must provide a “meaningful opportunity” for someone so designated to get notice of the factual basis for their detention and to contest their status, it is being less deferential.²⁷ And when a court says that a person *cannot* be detained (given the legislature’s current authorization) unless someone is charged with a crime, it is being *even less* deferential.

The same pattern can be observed when comparing, for example, deferential versions of the sort of “arbitrary and capricious” and “substantial evidence” review to the role of a typical court considering the veracity of a prosecutor’s charging document in the course of a criminal trial. A conviction generally requires convincing the court of an argument about the particular meaning of a statute, and proving facts establishing the

²³ For cases, *see supra* note 6.

²⁴ *Hamdi*, 542 U.S. at 510.

²⁵ I’m obviously assuming here that court decisions actually matter, and that (in many cases) they do have the potential to greatly affect how the executive branch does its work. That’s an assumption that makes considerable sense given how much time, money, and energy interest groups spend litigating, executive branch lawyers spend responding, and legislators spend writing laws that are supposed to affect the litigation process.

²⁶ *See Hamdi*, 316 F.3d 450, 473 (4th Cir. 2003).

²⁷ *See Hamdi*, 542 U.S. at 519.

presence of “elements” of the crime “beyond a reasonable doubt.”²⁸ Specific cases may involve different doctrinal bases and legal subtleties. But across this diverse doctrinal territory, one might imagine that a given point on a continuum (ranging from extreme deference to pronounced stringency) could reflect the extent of the hurdles government must overcome to prevail in the judicial arena. It also serves to determine the costs various players will bear once the review process has played out; less deference forces both the executive and the court to devote greater resources to resolving the relevant legal and factual questions. The important distinction here is not in the specific doctrines or procedures involved, but in the fact that under a standard of greater deference (whether determined by legislative enactments or subsequent judicial interpretation), the executive branch retains a greater chunk of power to decide how to use its discretion.

When deciding how much of that power to let executive authorities keep, courts and legislators make an implicit assumption across a bewildering array of contexts. They appear to make decisions premised on the assumption that a particular degree of stringency in review will apply, once articulated (and assuming it is actually followed) *across the board* to all similarly-situated cases.²⁹ This is perhaps partly a function of *stare decisis* and prevailing conceptions of horizontal equity, or perhaps even partly driven by inflated conceptions of judicial power to ensure that like cases are treated similarly. In fact courts treat horizontal equity as an important value, where deviations must be defended.³⁰ The same goes for virtually all the legislative mandates that courts implement. This renders troubling (in the eyes of many principled observers) the prospect of increasing the stringency of review without making the various costs prohibitive.

C. Audits Can Substitute For, or Supplement, Judicial Review

Now suppose court review isn’t available, that it is considered too costly for some reason, or that we want to supplement the allegedly laudatory impact of judicial review on executive behavior in some way. How else might legislators and the public police executive discretion? One answer can be found in what government organizations repeatedly do to the public: they audit. By thinking about alternative institutional arrangements, such as audits, we can get a better sense of the practical and political opportunities to nudge bureaucracies with discretionary legal powers away from failure. We can also clarify what our underlying goals are when discretion is reviewed, thereby trading off various possible ways of solving some of the institutional design problems

²⁸ See *In re Winship*, 397 U.S. 358 (1970).

²⁹ See, e.g., *Bush v. Lucas*, 462 U.S. 367, 388 (1983) (noting, in the context of reviewing an alleged retaliatory demotion and defamation claim against a federal official, that “[t]he costs associated with the review of disciplinary decisions are already significant – not only in monetary terms, but also in the time and energy of managerial personnel who must defend their decisions”). See also *Langevin v. Chenango Ct., Inc.* 447 F.2d 296 (2d Cir. 1971) (finding that judicial review of the Federal Housing Agency’s discretionary actions resulting in rent increase approvals in part on the basis that an unacceptably high number of rent increases would be subject to review). Here and in similar cases, the Court’s discussion of costs implicitly assumes that whatever costs are generated by the stringency of review the court adopts in the present case will be applied to all future cases with similar characteristics. See *supra* note 7 for examples of briefs making this argument.

³⁰ See *Cole*, *supra* note 1, at 2567.

associated with audits (like how courts should be involved, or whether audits should be publicly disclosed) that might be better or worse suited to achieving particular goals.

As I use the term here, an audit is a sustained, careful evaluation of a discrete decision drawn from a larger pool. In contrast to financial or more wide-ranging management audits, the audits of targeted discretion I discuss below treat each discretionary decision, like a decision to label a group as a specially designated terrorist organization, as the unit of analysis. Audits of targeted discretion would evaluate the information supporting the decision, its origins and reliability, contradictory information, and the broader context in which the decision took place. Though audits rarely take precisely the form I suggest here, the basic idea of using audits to learn what's going on in the world is neither mysterious nor rare.

Audits associated with taxation are among the most familiar. They take place in some form in most reliable tax collection systems, which, like our own, rely partially on self-reports.³¹ Many tax audits are not entirely random, which reduces their ability to provide a reliable picture of what's going on in the world. The less random the audits are, the less generalizable their results – and the easier it might be to evade them by avoiding the behaviors that raise the probability of being audited.³² From this perspective, one of the “purest” tax auditing programs in recent years (in the sense of being almost entirely random) was the Internal Revenue Service's Taxpayer Compliance Measurement Program (or TCMP). The following discussion emphasizes the tremendous informational value of such a program:

The last thorough tax gap study was for the year 1992, based on the 1988 TCMP. Noncompliance with individual and corporate income taxes was estimated to cost the Treasury about 18 percent of actual tax liability, which at 2002 levels of revenue would have amounted to \$223 billion. An average tax rate of 22 percent implies that there is about \$1 trillion of unreported income and illegitimate deductions.³³

Without TCMP audits, the federal government is unable to figure out the size of the “tax gap.” The program's cancellation has limited the government's ability to know how much is paid relative to what is owed, and who's particularly likely to be responsible for that gap.³⁴

Auditing-type activity also shows up in the pages of court opinions. Suppose Federal health care regulators and investigators suspect a health care clinic or nursing home of overcharging the federal government on Medicare payments. The government sends in investigators. Instead of figuring out the amount the clinic owes by reviewing each one of its files, investigators use audits to calculate the amount.³⁵ Courts reviewing

³¹ See generally JOEL SLEMROD AND JON BAKIJA, *TAXING OURSELVES: A CITIZEN'S GUIDE TO THE DEBATE OVER TAXES* (2004).

³² This last point is at the core of the explanation for why a “mixed strategy” is so valuable in the framework of game theory. See DAVID A. KREPS, *A COURSE IN MICROECONOMIC THEORY*, 381-83 (1993).

³³ Slemrod and Bakija, *supra* note 31, at 175.

³⁴ *Id.* (“[T]he estimates are based on data that is now over fifteen years old. But these are the best numbers around.”).

³⁵ See *Chaves Co. Home Health Serv., Inc. v. Sullivan*, 931 F.2d 914 (DC Cir. 1991).

this practice have repeatedly endorsed it, finding neither a conflict with the statute nor one with due process.³⁶ “Sample audits” also make an occasional appearance in class actions. When they do, courts (and litigants) confronted with an entire class of claims take a sample of those claims to get a better sense of what’s going on.³⁷

Something similar happens in the private sector, where “*your call may be monitored for quality assurance.*” The point is not just to prevent abuses or mistakes in discrete cases. It’s to learn more – as with tax collection -- about what’s really happening and how it can be improved. For example, insurance companies sometimes perform a process of “closed file review,” where they spend more money figuring out whether the amount of money paid out for a particular insurance claim was correctly calculated than they do paying out the claim itself.³⁸ It’s not hard to see why managers would rather know what their employers are doing. Nor is it surprising that random audits (at least when they happen with a sufficiently high probability) make it harder for the people or organizations being overseen to evade detection.³⁹

Not surprisingly, the federal government has created several bureaucracies capable of using audit-type techniques to investigate what government agencies actually do with their discretion. Very occasionally, government agencies audit the performance of their own workers.⁴⁰ The Government Accountability Office (originally the General Accounting Office, or GAO) was created early in the 20th century primarily to help Congress monitor the financial activities of the executive branch.⁴¹ In 1974, legislators gave the GAO power to review and analyze the implementation of government programs. Shortly thereafter, beginning in the middle of the 1970s and continuing over the next ten years or so, legislators began creating “Inspector General” offices in the federal government.⁴² Like the GAO, the Inspectors General have the legal power to investigate how federal officials use their targeted discretion. The existence of these structures indicates the potentially important role that audits can play in shaping how the federal government uses its targeted discretion. Whether these bureaucracies actually perform such audits is another matter, to which I return below.

To understand how audits would work, imagine a world much like our own, where some decisions (like environmental regulatory rules, voting rights regulations, and criminal convictions) are subject to stringent judicial review. Other executive branch decisions (like special designation as a terrorist organization or as an enemy combatant) are subject to less stringent review. For fairly obvious reasons, we (or if you prefer the parlance of game theory, the “principals”) want to know how the government is using its

³⁶ *Id.*

³⁷ See, e.g., *McComber v. Traveles Property & Casualty*, Case No. X03CV990496761S, Superior Ct. of Conn. (Jud. Distr. Of New Britain)(May 26, 2004)(unpublished disposition).

³⁸ See J. David Cummins and Sharon Tennyson, *Controlling Automobile Insurance Costs*, 6 J. ECON. PERSP. 95 (1992).

³⁹ See Kreps, *supra* note 32, at 763-64.

⁴⁰ See Exec. Order No. 12,564, 51 FED. REG. 32889 (Sept. 15, 1986)(subjecting numerous categories of federal employees to random drug tests).

⁴¹ See John T. Rourke, *The GAO: An Evolving Role*, 38 PUB. ADMIN. REV. 453 (1978).

⁴² See PAUL LIGHT, *MONITORING GOVERNMENT: INSPECTORS GENERAL AND THE SEARCH FOR ACCOUNTABILITY* 42 (1993).

discretion. The problem is we have a limited budget to review decisions, and we are concerned about over-detering the executive branch. Earlier I noted that a key feature of judicial review is that courts and legislators tend to pick a standard of deference that's supposed to apply to all cases in a particular class. What audits do is to introduce an alternative means of review that allows for variation in *both* the standard of deference used to review cases *as well as* the number of cases actually reviewed. In exchange for reviewing fewer cases, whoever is conducting the audits can demand more evidence from the executive branch, more justification, and more access to information – all at a lower cost than what would be incurred if the *same* standard of deference applied to every decision.⁴³

The process would unfold along the following lines. First, an auditor would define some discrete set of targeted decisions to analyze (i.e., all summary exclusions at the border, all enemy combatant designations, all discretionary – as opposed to mandatory – container inspections, or all decisions to freeze assets). Second, the auditor would randomly choose some number or percentage of decisions to audit. Third, those decisions, would be reviewed far more stringently than a court would review the *whole class* of decisions. If a court (as with border inspection decisions) provides almost no review, the auditor would gather all available information about how the decision took place, what its effect was, what the secondary inspector knew when he denied entry, and what other agencies know that might be relevant to the decision. If a court reviews IEEPA designations under a highly permissive version of the “arbitrary and capricious” and “substantial evidence” tests, the auditor would instead gather information on *how* an administrative record was compiled – not just on what it purports to say. In doing this, the auditor would apply some kind of standard (which I discuss below) either drawn from the purposes of the statutes in question, or perhaps even based on what the executive branch says it is trying to accomplish through its actions. Fourth, the results of the audit would be made available to legislators and the public, which may (under certain conditions) then pressure the agency to make modifications in its conduct. Intelligence information could be used in these determinations because it can be reviewed *in camera*.⁴⁴ Obviously, a lot in this scenario depends on whether legislators and the public react to the audits. While they might sometimes ignore those results, the media's reaction to GAO and Inspector General reports suggests that these audits could be quite salient.⁴⁵ Judicial review would continue in the background at whatever standard of deference courts and legislatures choose – though it's certainly plausible to think that some courts might approach their cases differently as a result of what the audits revealed. They might even evaluate executive clamoring for deference by weighing whether a reliable audit

⁴³ While this proposal is reminiscent of Mashaw's call for a super-agency to review benefit determinations, in this Part I try to offer a more detailed case for audits than has been offered before. I also paint a picture of how they would work, and why judicial review as it currently functions leaves such a striking gap in our ability to review some kinds of executive discretionary decisions. Cf. JERRY L. MASHAW ET AL., SOCIAL SECURITY HEARINGS AND APPEALS: A STUDY OF THE SOCIAL SECURITY HEARING SYSTEM (1978).

⁴⁴ Cf. *United States v. Isa*, 923 F.2d 1300 (8th Cir. 1991)(using *in camera* review of information obtained by federal agents through wiretaps authorized under the Foreign Intelligence Surveillance Act and denying aggrieved party's motion for suppression); JOHN HART ELY, WAR AND RESPONSIBILITY (1993)(final chapter – discussing how fear of leaking is overblown).

⁴⁵ See *infra* Part IIIa.

system is in place. Although audits would not necessarily provide relief to every aggrieved person or group (that depends on how they're designed), they would help legislators, organized interest groups, and the public to learn far more about what government does than is currently known.⁴⁶

A simple model serves to demonstrate the potential differences between audits and traditional judicial review. Models obviously abstract from the real world. While they should be used cautiously, my aim in using one here is simply to offer a more concrete illustration of the critical ideas. Begin by accepting the executive branch's premise that more intrusive judicial review is a problem because it is *too* costly, with cost here referring to the whole gamut of allegedly adverse consequences associated with greater review stringency. This suggests (plausibly enough) that there is a relationship between the stringency of judicial review and its cost. The cost of review reflects a couple of things. We should expect that the more cases reviewed, the higher is the direct cost incurred by the court (or some other reviewing authority) when examining cases, and by the executive branch when providing information and defending its actions on an individual case. Moreover, the more cases reviewed (or at least eligible for review) out of a total pool of cases, the more that the benefits of targeted discretion might dissipate. Decisions may be slower (either because of the resources consumed by the review process, or simply because the fact they will be reviewed leads the executive branch to make the initial determination more judiciously). The risk of over-deterring may also be greater as the proportion of cases eligible for review rises towards 100%. On the other hand, more stringent judicial review is valuable because it is more likely to reveal problems in targeted discretionary decisions.

A few plausible assumptions illustrate the situation. Let the terms C_d and C_i denote direct and indirect costs, respectively, of reviewing cases with a particular degree of stringency. Let $Pr(\pi)$ denote the probability that a problem is discovered given a particular degree of stringency in the review process, which is represented by an increasing parameter S that begins at zero and increases to 1, which represents maximally intrusive review. Let N_c denote the proportion of cases reviewed in a given class of cases. Here's how all this fits together:

- C_d and C_i increase as a linear function of S , such that $C=C_d + C_i$ and $C=f(S)$, such that one can define a two dimensional space consisting of a cost dimension C running vertically, and a perpendicular dimension stringency of review dimension S running horizontally.
- For any given set of cases and stringency of review, both C_d and C_i increase as N_c increases. This implies that the slope of the line defining the relationship between S and C becomes more elastic as the proportion of cases reviewed *decreases*.
- $Pr(\pi)$ increases as a linear function of S , such that one can *also* define a two dimensional space consisting of a probability of discovering problems dimension $Pr(\pi)$, and a precisely orthogonal stringency of review dimension S .

⁴⁶ Depending on the assumptions made about the political system, legislators and the public might respond to the audits in ways that *would* provide relief to all or some of the people aggrieved by problematic applications of targeted executive discretion.

- At some levels of stringency, the probability of discovering problems falls below an acceptable minimum (where “acceptable minimum” means the point below which, according to a social consensus, the probability of discovering problems should not fall). There is a unique point on the S dimension (call it S_{min}), indicating the minimum acceptable degree of stringency.
- For any given N_c , there is some point C_{max} representing the maximum cost of review that society is willing to bear in reviewing targeted discretion. Because C increases as a function S , C_{max} corresponds to a maximum degree of stringency of review (call this S_{max}).

Figure 1 shows how these parameters interact in situations where judicial review is likely to work reasonably well. Any point on the stringency axis (at the top and bottom of the figure) would have corresponding points, indicating the cost of that degree of stringency and its associated probability of discovering problems, on the *two* perpendicular axes. The upper part of the figure shows the relationship between the costs of review and stringency for a given proportion of cases reviewed. The cost-stringency line has roughly the same slope as the probability-stringency line in the lower figure, which indicates the relationship between stringency and the probability of discovering problems.⁴⁷ The space between the lowest acceptable bound of review (S_{min}) and the highest permissible cost (which corresponds to S_{max}) is what I call the “feasible review set.” This is the space where courts (or legislators) have some flexibility in setting the stringency of the doctrine. Any level of stringency lower than S_{min} means there won’t be a high enough probability of discovering problems, and a level higher than S_{max} effectively breaks the bank – either because the review process itself becomes too expensive or because of there is too much interference with the benefits of discretion in the executive branch.⁴⁸

[FIGURE 1 HERE]

Switch now from thinking about judicial review to thinking about audits, and a subtle but important difference emerges in the situation. Because we have assumed that the relationship between C and S becomes more elastic as the proportion of cases reviewed *decreases*, then audits allow a higher degree of stringency of review (and thereby a lower standard of deference) for the *same* cost. Figure 2 shows how this looks. In contrast with Figure 1, the slope of the cost-stringency line is now flatter than that of the probability-stringency line in the lower portion of the figure.⁴⁹ Which means the feasible review set suddenly expands dramatically while the costs of review remain the same. The implication is that, in exchange for lower costs by reviewing a smaller proportion of cases, the auditor gains the chance to review cases more thoroughly – to ask

⁴⁷ Thus, in this first figure $C=S=Pr(\pi)$.

⁴⁸ I’m assuming here that S_{min} and S_{max} are set exogenously, but a more complex model might derive these endogenously as responses to what is learned over time. For present purposes, we might assume that S_{max} reflects some welfarist conception of a budget constraint as well as a technical judgment of how much executive branch discretion is desirable for a particular set of decisions. Meanwhile, S_{min} might reflect the minimal degree of stringency necessary to force the executive branch (or one of its organizational units) to behave appropriately.

⁴⁹ In contrast with the previous figure, the relationship is now roughly $C=3/2 S = 3/2 Pr(\pi)$.

for more evidence, to require a more explicit showing of facts in a record, to inquire into how a particular record was developed, and ultimately, to better understand what the pressures may have affected the decisionmaker and distorted her decision. What both Figures 1 and 2 have in common, nonetheless, is that the decisionmaker (whether a reviewing court or an auditor) could choose a level of stringency above S_{min} – or that level of stringency which is associated with the minimum acceptable probability of discovering problems.

[FIGURE 2 HERE]

The preceding discussion shows two differences between audits and traditional forms of judicial review. With audits the costs of audits can be adjusted by changing the number of cases reviews, and not just the stringency of that review. In addition, audits would not reflect the distortions imposed by the litigation process. The cases reviewed would not be a function of who can obtain representation or what particular fact pattern entices judicial attention. As I explain below, both of these characteristics turn out to be important in making the case for supplementing judicial review with audits of executive discretion.

D. Traditional Judicial Review Suffers From Limitations

So far I have shown how audits can substitute for judicial review, but I have not yet addressed why it would be preferable (from what might be broadly termed a social welfarist perspective) to live in a world where judicial review is supplemented by audit review. To understand judicial review's merits as a way of overseeing executive discretion, we must review some of the costs and benefits of deviating from the conventional limited-discretion baseline discussed above. With this review and the aid of some specific examples involving considerable degrees of executive discretion, we can better understand what audits can contribute.

Given the allegedly intimate link between accountability and court review, any effort to depart from a baseline of stringent review should be continued on a satisfactory accounting of the benefit associated with departing from the limited-discretion baseline? A fairly obvious one is speed, or what is commonly termed “efficiency.” Some decisions need to be made quickly if they're going to matter. Suppose policymakers confront a possible outbreak of avian flu virus. They may consider imposing a quarantine. They must decide quickly whether American airports will receive flights from the affected country. To delay the decision effectively becomes a decision to let the planes land. Even if it is possible to wait, it may cost a lot to do so. The Treasury can wait to freeze a suspicious charity's assets, but those assets may soon leave the group's coffers for some sleepless island bank secrecy haven. Letting executive authorities have discretion lets them not only decide quickly – and the saved time can translate into money, extra safety, and convenience. The point is not lost on courts reviewing many of the federal government's national security decisions.⁵⁰ Nor is it lost on courts and scholars writing about other aspects of public law – such as those concerned about the “ossification” of regulatory rules.⁵¹ On a related note, less review also saves two kinds of resources: those

⁵⁰ See *Water Keeper Alliance v. U.S. Dept. of Defense*, 271 F.3d 21 (1st Cir. 2001).

⁵¹ See *Cornelius Kerwin, Rulemaking* (3^d ed. 1999). See also McGarity, *supra* note 22, at 21-23.

the court or other reviewing authority would expend on analyzing a case, and those that the government would spend defending itself. These costs are likely to be especially salient because courts, relying on some version of *stare decisis* or horizontal equity norms, assume they're fashioning a standard that will apply to all (or nearly all) similar cases.⁵²

Another argument for letting government have a relatively free rein in applying discretion is rooted in the allegedly superior technical competence of the executive branch.⁵³ The conventional wisdom is that agencies and the executive branch have greater expertise than the judges who might review their decisions. This point shows up time and again in judicial decisions in a wide range of domains, from environmental policy to police investigations.⁵⁴ No doubt that expertise is valuable. The more some reviewing authority intervenes, the greater the risk that expert decisions will be undone. More intervention may even dilute the incentives of decisionmakers to develop and use expertise.⁵⁵

Discretion may also have a role in helping government harmonize competing goals, trading off some desired goals against further delays (for example) in achieving less compelling policy objectives. Providing more inspectors for port security might require a somewhat modified inspection regime in the Northern Border to make up for the more limited personnel there. If OSHA assigns some of its best inspectors to implement a new regulation instead of enforcing an old one, it may have to change the threshold for deciding to inspect a plant in order to make up for the change in the quality of the officials enforcing the old regulation. Examples like these emphasize how discretionary decisions do not happen in a vacuum. The president and his advisors may decide to devote fewer resources to freezing terrorist assets if they feel there's been a n improvement in their ability to disrupt the travel plans of suspected terrorists. Regulators with less resources for reviewing disclosure filings may feel the need to conduct more inspections. Discretion lets executive authorities harmonize decisions in different policy domains. Other things being equal, it seems plausible that greater power for courts (or some other reviewing authority) to intervene and undermine executive decisions would make it harder for the executive branch to harmonize decisions effectively.⁵⁶

Finally, courts and commentators sometimes talk about the value of making the executive branch "accountable" for its decisions. In an ironic twist, the rhetoric of accountability that so often bolsters arguments for stringent judicial review sometimes serves precisely the opposite goal. The argument goes something like this. The less that court (or other external) intervention encroaches on the executive's domain, the more that legislators, organized interest groups and the larger public can focus on rewarding or

⁵² See *supra* note 29.

⁵³ See *Perez v. F.B.I.*, 71 F.3d 513 (5th Cir. 1995). *But see* *North Jersey Media Group, Inc. v. Ashcroft*, 308 F.3d 198 (3rd Cir. 2002).

⁵⁴ See *Vermont Yankee Nuclear Power Corp. v. NRDC*, 439 U.S. 961 (1978); *I.N.S. v. Lopez-Mendoza*, 468 U.S. 1032 (1984).

⁵⁵ Cf. KEITH KREHBIEL, *INFORMATION AND LEGISLATIVE ORGANIZATION* (1991).

⁵⁶ See *Greenbaum v. E.P.A.*, 370 F.3d 527 (6th Cir. 2004).

punishing the executive (or the inferior officer) for her decisions.⁵⁷ This position implies not only a reluctance to see courts throw sand in the gears of some hypothetical scheme for accountability, but a confidence that an accounting will indeed be rendered to either superior officers or the public. Thus courts observe (as did this one in declining to engage in review of prosecutorial discretion) that “while this discretion is subject to abuse or misuse just as is judicial discretion, deviations from [the prosecutor’s] duty as an agent of the Executive are to be dealt with by his superiors.”⁵⁸

These arguments have each have a grain of truth. What they don’t address is how much – and how easily – discretion can be abused, whether the context is social security benefit payments, border screening, enemy combatant designations, or prosecutorial enforcement. Consider, for example, what could be called the “learning costs” problem. Executive branch bureaucracies and the people who work in them spend their days (ostensibly) carrying out legal mandates. People who work there do that in part by relying on expertise. They hone that expertise by learning from their environment, and correcting their mistakes. But if no external authority monitors the bureaucracy, then those who work there may be unwilling or unable to learn much of anything. In fact, several scholars have suggested that external court review helps bureaucratic institutions learn. That belief is not always fully theorized – and of course court review carries with it the risks of over-detering executive branch activity.⁵⁹ No doubt sometimes an inspector’s good conscience or an agency’s strong internal culture contribute to reasonable decisions about what assets to freeze or who should be labeled an enemy combatant. Nonetheless, it is certainly plausible to assume that such desirable circumstances do not always transpire, and that judicial review helps create conditions that foster learning. I make this claim for two separate reasons. First, a substantial body of research suggests that people learn when they have reason to do so.⁶⁰ Other things being equal, the dilution of review may deprive individuals of reasons to learn (or, at least, to learn with the same intensity than they would if review were more stringent). This assumes, quite plausibly, that a review process turning up mistakes can be embarrassing to people, or that people in the agency may otherwise suffer some costs if they face some kind of review process that does not go well.⁶¹ Second, organizations develop routines that blind them. As Diane Vaughan wrote in her study of the *Challenger* launch decision:

Possibly the most significant lesson from the *Challenger* case is how environmental and organizational contingencies create prerational forces that shape worldview, normalizing signals of potential danger, resulting in mistakes

⁵⁷ See, e.g., *Chevron v. NRDC*, 467 U.S. 837, 865 (1984) (“While agencies are not directly accountable to the people, the Chief Executive is...”); *Motor Vehicle Mfrs. Ass’n of U.S. v. State Farm*, 463 U.S. 29 (1983) (Rehnquist, J., concurring in part and dissenting in part).

⁵⁸ *Newman v. United States*, 382 F.2d 479 (D.C. Cir. 1967).

⁵⁹ This last point would seem to depend crucially on implicit assumptions about the extent to which the players involved would find it valuable to avoid having their actions vacated by a court or otherwise found invalid.

⁶⁰ See ARTHUR LUPA AND MATHEW MCCUBBINS, *THE DEMOCRATIC DILEMMA: CAN CITIZENS LEARN WHAT THEY NEED TO KNOW?* (2001).

⁶¹ This is certainly true in the case of people who work in offices whose broad performance is review by Inspectors General or the GAO. See *infra* notes 124 and 125.

with harmful human consequences. The explanation of the *Challenger* launch is a story of how people who worked together developed patterns that blinded them to the consequences of their actions. It is not only about the development of norms but about the incremental expansion of normative boundaries: how small changes – new behaviors that were slight deviations from the normal course of events – gradually became the norm, providing a basis for accepting additional deviance. *No rules were violated; there was no intent to do harm. Yet harm was done.*⁶² [Emphasis added].

External review may elucidate things that people inside the organization fail to appreciate. Outsiders may see things not despite, but precisely because of, the absence of expertise. Which means that even if discretion plays a vital role in creating the incentives for people to gather expertise and for other reasons discussed previously, its abundance may diminish opportunities for learning from mistaken enemy combatant designations, border inspection decisions, asset freezing determinations, and health or safety inspections. The larger point is that the most attractive kinds of organizational learning – where the organization learns to achieve important goals better and more efficiently – is likely to be rarely encountered, if in fact it is encountered at all.⁶³ Watering down or forgoing judicial review altogether leaves the problem of how agencies will learn from their mistakes, and indeed, how agencies will even realize that they have made a mistake.⁶⁴

Large grants of discretion can have at least two other problematic consequences. In some cases, executive branch officials may succumb to the temptation to use their discretion to create an appealing impression among the public. I discuss this problem at greater length elsewhere,⁶⁵ but the basic insight is a simple one. Discretionary actions can serve as a sort of signal that the public (or political superiors) can use in forming judgments about the competence of the executive branch (or an organization within it). As long as the public's impressions of the executive branch's expertise, success, ability, and resolve are influenced in part by discretionary actions, then those actions will become tempting levers to create favorable public perceptions. Frozen assets and specially-designated terrorist organizations send the message that the executive branch knows what it's doing when it may not. This state of affairs may skew citizens' ability to evaluate the

⁶² DIANE VAUGHAN, *THE CHALLENGER LAUNCH DECISION: RISKY TECHNOLOGY, CULTURE, AND DEVIANCE AT NASA* 409-410 (1996)

⁶³ See Michal Tamuz, *Learning Disabilities for Regulators: The Perils of Organizational Learning in the Air Transportation Industry*, 33 ADMIN. & SOC. 276 (2001); James G. March et al., *Learning From Samples of One or Fewer*, 2 ORGANIZ. SCI. 1 (1991).

⁶⁴ See Mashaw and Harfst, *supra* note 22, for a discussion of how a regulatory agency (in that case, NHTSA) learned to use alternative policymaking strategies to avoid the costs associated with judicial review. The example serves to emphasize two points, both of which are relevant to the present discussion: (a) that agencies appear to react to judicial review, and (b) that organizational learning is not necessarily associated with learning too achieve the most valuable organizational goals better. Learning can be a bad thing; no doubt that organizations making large numbers of discretionary decisions that rarely if ever get reviewed (or, if reviewed, rarely get scrutinized carefully) probably learn that they can shift resources, time, attention, quality control, and strict adherence to legal or aspirational goals away from those decisions and towards other pursuits. The question is how to encourage the most desirable kinds of learning.

⁶⁵ See Mariano-Florentino Cuéllar, *The Mismatch Between State Power and State Capacity in Transnational Law Enforcement*, 22 BERKELEY J. INT'L L. 15 (2004).

effectiveness of their own government. And the discretionary actions may themselves have costs, including the creation of perverse incentives for regulated groups,⁶⁶ diminished compliance with treaties, or simply the individual mistreatment suffered by individual detainees (for example) whose weeks as enemy combatants became months and then years before ending in freedom. Finally, just as discretion allows political authorities to engage in subtle, politically-motivated self-dealing, it can also lead to some employees engaging in blatant, willful malfeasance.⁶⁷

No one in their right mind who reflects for more than a moment on these costs could possibly deny that they must be weighed against discretion's benefits. Without accepting the need to balance (in principle), discussions of discretion in courts and scholarly circles become incoherent. Massive discretion cannot be defended in the abstract because it is neither inherently good or bad. Its consequences depend on the substance of the decision, its context, and the complex interplay of incentives and capacity affecting the decision-maker. Should prosecutors be able to charge whomever they choose without looking over their shoulder? It depends. Should the president retain the power to control who's detained as an enemy combatant or what assets, of which charity, get frozen? All these powers are granted not because discretion's inherently valuable but because courts and legislatures claim good things will come as a result, so that the costs of discretion are worth bearing. So airport screeners keep choosing passengers for secondary inspection, welfare offices keep paying benefits or denying applications for them, and prosecutors keep choosing whether to charge a suspect.

Which brings us to the question of how to strike that balance between costs and benefits in the specific contexts that together constitute a messy, tangled world of politics, institutional pressures, economic constraints, ambitious legal mandates and uncertainty. Consider some examples.

E. Examples

i. Freezing Financial Assets of Suspected Terrorist Supporters

Governments try to freeze the assets of groups that threaten their national security. If a charity is raising funds for terrorists, the U.S. government will want to freeze those assets. Doing so reduces the chance that those assets will fund the lives of terrorists but it also works as a sort of punishment against people or groups who should have known better.

The law gives government officials some powerful tools to do this. One of them, designations under the International Emergency Economic Powers Act (or "IEEPA"),⁶⁸ lets the president "block," or freeze access to, any property subject to United States jurisdiction, when two conditions apply. First, the property in question must be

⁶⁶ *See id.*

⁶⁷ *See* THE ABU GHARIB INVESTIGATIONS: THE OFFICIAL REPORT OF THE INDEPENDENT PANEL AND PENTAGON ON THE SHOCKING PRISONER ABUSE IN IRAQ (2005).

⁶⁸ International Emergency Economic Powers Act, 50 U.S.C.A. § 1701 et seq ("IEEPA"). *See generally* James J. Savage, *Executive Use of the International Emergency Economic Powers Act – Evolution Through the Terrorist and Taliban Sanctions*, 10 CURRENTS: INT'L TRADE L.J. 28 (2001).

something in which a foreign country or national has an interest.⁶⁹ This constraint turns out to be not much of a limitation on the president's power, since courts have found that the "foreign" interest does not have to be a legal interest of any kind. The mere fact that an organization has foreign beneficiaries may be enough, in fact, for a court to say that it has a "foreign interest."⁷⁰ Second, the president must use this power only during an emergency.⁷¹ That's not much of a limitation, either. The emergency must involve an "unusual and extraordinary threat" giving rise to the emergency must have its source partly outside the United States.⁷² It might pose a threat to the "national security, foreign policy, or economy of the United States."⁷³ The combined effect of this expansive language and traditional judicial deference on matters of national security and foreign affairs have made it relatively easy for presidents to declare a number of emergencies (about ten of which are currently in effect).⁷⁴ Courts have yet to find that, under the terms of IEEPA, a supposed emergency does not, in fact, exist.

In a series of executive orders, the President's delegated authority under IEEPA to the Secretaries of State and Treasury. Under the resulting system, the Treasury's Office of Foreign Asset Control blocks the assets of groups that branded "Specially Designated Terrorist Organizations."⁷⁵ Once a group becomes a specially designated terrorist organization, it loses control over its fate. As a lawyer for a designated group recently wrote, the organization is likely to shut down and lay off its employees:

Acting under the blocking notice, government agents entered HLF's offices in Texas, New Jersey, Illinois, and California, seized and removed the contents (including everything from documents to office equipment to employees' personal effects), and froze HLF's bank accounts, which contained millions of dollars in charitable contributions. As of that day, HLF ceased operations. Its employees lost their jobs.⁷⁶

The impact of OFAC's orders is to block the organization's funds, regardless of where in the financial system they happen to be. The designation also means that people providing "material support" to the group commit a serious federal crime.⁷⁷

Take a closer look at how the court reviews the government's designations. In the recent *Holy Land Foundation* case, the State and Treasury departments used IEEPA powers delegated to them by the president.⁷⁸ Agency staff issued a blocking order. The blockage limits what the charity can do. Its powers and options are drastically constrained as a result of the blocking order. When the court considers whether a

⁶⁹ *Id.*

⁷⁰ See *Global Relief Fdn. v. O'Neill*, 315 F.3d 748 (7th Cir. 2002).

⁷¹ See IEEPA, *supra* note 68.

⁷² See *Holy Land Fdn. v. Ashcroft*, 333 F.3d 156, 159 (D.C. Cir. 2003).

⁷³ *Id.*

⁷⁴ See generally Jason Luong, Note, *Forcing Constraint: The Case for Amending the International Emergency Economic Powers Act*, 78 TEX. L. REV. 1181 (2000).

⁷⁵ 333 F.3d at 160.

⁷⁶ See John D. Cline, *The President's Power to Seize Property in the Post-September 11 World: The International Emergency Economic Powers Act*, CHAMPION (Sept/Oct. 2003).

⁷⁷ See *Humanitarian Law Project v. Reno*, 205 F.3d 1130 (9th Cir. 2000).

⁷⁸ 333 F.3d at 159-60.

designation was legal, it defines a certain standard of review to circumscribe its inquiry. Specifically, it considers whether the decision is “arbitrary and capricious” under the terms of the Administrative Procedure Act. But it interprets IEEPA and the national security context in which these decisions take place to require a highly deferential form of review. Now the District Court took that standard and seemed to push it about as far as it would go. The key factual question, the District Court and the litigants agreed, was the extent of HLF’s connection to Hamas, another specially designated terrorist organization (and one that, at least at this point, few people had reason to doubt as a “terrorist organization”). The district court conducted a careful examination of the record and uncovered “ample” evidence that:

- (1) HLF has had financial connects to Hamas since its creation in 1989; (2) HLF leaders have been actively involved in various meetings with Hamas leaders; (3) HLF funds Hamas-controlled charitable organizations; (4) HLF provides financial support to the orphans and families of Hamas martyrs and prisoners; (5) HLF’s Jerusalem office acted on behalf of Hamas; and (6) FBI informants reliably reported that HLF funds Hamas.⁷⁹

This kind of review might strike some observers as particularly thorough. That perception is mistaken, for at least two reasons. First, the court considers only whether the decision was “arbitrary and capricious,” and based on “substantial evidence,” not whether it was right or wrong. That determination, moreover, reflects a statutory text (the APA and IEEPA) and tradition that makes the court’s inquiry extremely deferential and perhaps helps explain why so few of these determinations get challenged in court (because it’s not clear what will be gained). Second, as a practical matter, the court’s inquiry (even where, as in *Holy Land*, the district court pushes the envelope in terms of the stringency of its review) begins and ends with the record that the government itself compiles. As the district court itself noted in this case, the arbitrary and capricious standard “does not allow the courts to undertake their own fact-finding, but to review the agency’s record to determine whether the agency’s decision was supported by a rational basis.”⁸⁰ That record may be a tremendously accurate complication of the government’s evidence. Or it may be patently misleading. Nothing requires that the government report evidence tending to cast doubt on its contentions.⁸¹ Nor does the court interview the sources on which the record is based. Which means that, even in the best of circumstances, the court’s review is only as good as the record.

The flip side of this point is that court review is likely to exert only a limited impact on the quality of that record. If that record turns out to be an empty folder, then the court would almost certainly vacate the designation without ever getting to the arbitrary and capricious test, because of the absence of “substantial evidence.” On the other hand, officials who want to evade that possibility need only make sure there is a thick enough record to make it hard for the court to conclude that such a record makes the

⁷⁹ 333 F.3d at 161.

⁸⁰ See *Holy Land*, 219 F.Supp.2d at 67.

⁸¹ Ironically enough, the *Holy Land* appeals panel suggested that the government’s position was strengthened by the fact that “there was no plausible evidence presented which showed that [ties to Hamas] had been severed.” 333 F.3d at 162.

designation look totally arbitrary. The thing is, the record itself is based on decisions that are essentially immune from review. Which makes them prone to the familiar political pressures and distortions that permeate government bureaucracy.⁸²

ii. Government-Run “No Fly” Lists

When passengers check in at American airports, their names are compared to those on a list provided to airlines by the Transportation Security Administration (TSA).⁸³ Predictably enough, the point of the list is to thwart the (travel) plans of those who might prove dangerous on board a flight. A subsidiary purpose may also be to promote the questioning and apprehension of suspicious individuals who may be sought by law enforcement authorities. Although the aforementioned list (technically an element of a system known as CAPPs I) is considered by TSA to be an important component of aviation security, recent efforts have focused on supplementing it with a more elaborate safeguard.⁸⁴ The agency has been rushing to implement a long-delayed new system, since shortly after September 11, but the process has proven fraught with delays.⁸⁵ The advent of the new system, however, is unlikely to lead to the complete demise of some version of the current process, as the government will likely retain a “core” list of people who should be detained when they attempt to travel.⁸⁶

In practice, TSA and the airlines have run into problems using the list. It is not applied consistently. Sometimes people whose names are merely similar to those on the list are detained. It appears as though even individuals who are actually named on the list are virtually never completely denied the chance to fly; they are instead detained, at times forced to miss their flight, and subjected to extensive questioning. Even supposing assuming some of the problems with the list’s use could be remedied, it would seem as though there should be some procedure to police the list, so that people who are erroneously placed there could be taken off.

Nothing of the sort exists.⁸⁷ Instead, people who plainly should not be on the list (including, for example, a member of the Air Force reporting for duty) are bewildered to learn that they cannot even find out *how* their name appeared on the list, let alone what

⁸² Note that I am not suggesting that the manipulation of the record need be conscious or explicit. A number of pressures and considerable number of investigators, analysts, spies, lawyers, and higher level officials whose work influences the record that the court reviews. As long as they feel at least some subtle pressure to support the conclusion that a designation should be made, they may fail to consider countervailing arguments, or the potential consequences of an “erroneous” designation (i.e., erroneous in the sense of not complying with the statute, the president’s executive order, the “arbitrary and capricious” standard, or the executive branch’s stated goals for using the IEEPA emergency powers)].

⁸³ See Government Accountability Office, *Aviation Security: Computer-Assisted Passenger Prescreening System Faces Significant Implementation Challenges*, GAO 04-385 (Feb. 2004)(“GAO Report”).

⁸⁴ See generally *Green v. TSA*, 351 F.Supp.2d 1119 (W.D. Wash. 2005).

⁸⁵ See GAO Report, *supra* note 83, at 9. For developments since February 2004, see Government Accountability Office, *Aviation Security: Secure Flight Development and Testing Under Way, but Risks Should Be Managed as System is Further Developed*, GAO-05-356 (March 2005).

⁸⁶ See *EPIC v. Dep’t of Homeland Security*, ___ F.Supp.2d ___, 2005 WL 1745303 (July 25, 2005).

⁸⁷ See GAO Report, *supra* note 83, at 13 (discussing the absence of internal verification systems in TSA-administered passenger prescreening programs); *id.* at 36-38 (indicating how GAO did not itself audit names placed on the list as part of its methodology).

must be done to remove it.⁸⁸ TSA may be understandably reluctant to reveal the quality of its methods and sources, either because it seeks to avoid sensitizing potential terrorists to the extent of their strengths or alerting the larger public to the extent of their weaknesses. But that still leaves the question of how the quality of the list will be policed. The question becomes more pressing given the apparent absence of any reliable, consistent review mechanism to ensure the names on the list belong there.⁸⁹ Reasonable people can differ with respect to how much discretion TSA *should* have in deciding what information to evaluate when deciding whether to place someone on the list, and whether (on the basis of that information) a name belongs there. What's harder to dispute is the absence of any statutory provision for administrative or judicial mechanisms to resolve the problem. This leaves aggrieved parties with the option of a lawsuit claiming that their treatment as a result of being placed on the list amounts to a due process violation. Such a lawsuit is currently pending. But regardless of the outcome, the no-fly list problem illustrates why due process claims may be a poor vehicle for policing this type of discretion. Assuming the litigants persuade a court that the problems they confront while flying amount to an interference with a protected liberty or (less likely) property interest, they would still have to persuade a court that existing procedures violate the interest-balancing test rooted in *Mathews v. Eldrige*.⁹⁰ That would be a tough sell given the strength of the interest the government would assert in promoting aviation security. Which leaves the quality of the list dependent on the bureaucracy's behavior.

iii. Launching Major Occupational and Environmental Safety Inspections

In criminal justice, people and organizations tend to face severe penalties only after they're convicted or admit their guilt. When legislators approve statutes creating major regulatory programs, they often build in a different mechanism. Implicitly (and sometimes explicitly), their enactments recognize that using the criminal justice system to punish violations is cumbersome. Instead, many laws allow regulators to levy fines or issue orders restricting certain activities with more limited (if any) court intervention. Although they vary in the relevant legal standard or the size of the maximum fine, those orders have an effect before judicial intervention. Even after that intervention, it's not at all clear how well the stringency of review provided by courts (which tends to conform to some variation of the "arbitrary and capricious" or "abuse of discretion" standards) strikes the most desirable balance between restraining abuse and providing regulatory flexibility.

Consider two examples. When the Occupational Safety and Health Administration (OSHA) has some reason to think businesses are violating their general duty to provide a safe working environment, it can issue abatement orders and citations.⁹¹ Parties who fail to abate as the agency requires face additional penalties, as do those who

⁸⁸ Some individuals apparently on the list who write their congressional representative have been told to bring along multiple forms of identification to prove they are not the person named on the list. This obviously does not address the problem of the person *actually* named on the list who seeks to contest that designation.

⁸⁹ See GAO Report, *supra* note 83, at 13.

⁹⁰ See Green, 351 F.Supp 2d. at 1124-7 (denying due process challenge to the no-fly list from passengers who had been detained but eventually allowed to fly).

⁹¹ See 29 U.S.C. §§664(a), 666(a), 666(b), and 666(c).

engage in “serious” or willful” violations (even before an abatement order is issued).⁹² Although parties may (and often do) contest citations, doing so is expensive, which means some parties just pay the relatively meager fines OSHA tends to assess instead of contesting them. Penalties for violating compliance orders are considerably more severe under environmental statutes, like the Clean Air Act.⁹³ Under that statute, the Environmental Protection Agency (EPA) can issue an Administrative Compliance Order (ACO) on the basis of “any information available to it,” directing a regulated party – such as an electricity generation plant – or state agency to comply with the Clean Air Act’s requirements. While the ACO doesn’t allow EPA to impose finds or other penalties directly, the order triggers provisions imposing civil or criminal penalties for violation of the order.⁹⁴ Under the terms of the Act, it initially appeared as though judicial review of an ACO was supposed to focus on whether the regulated party violated the terms of the order, not whether the EPA was right to issue it in the first place.⁹⁵ This has understandably raised questions about how the order itself should be reviewed. In recent cases, the Supreme Court and several circuit courts have left some uncertainty about whether the ACO structure withstands constitutional scrutiny given its due process implications. At least one circuit court found ACOs not to be final agency actions (which is required for circuit court review in these cases), thereby rendering them unreviewable and raising the due process problem.⁹⁶ The Supreme Court declined to review this case, and instead – in a separate case – upheld a Ninth Circuit opinion holding ACOs to be final agency actions and reviewing them under the “arbitrary and capricious” standard.⁹⁷

In both cases, the relevant agencies retain considerable discretion in imposing compliance orders. OSHA obviously has it when it issues citations and abatement orders, some of which are not challenged subsequently. The environmental orders must survive arbitrary and capricious review if challenged (in circuits that didn’t find the whole structure unconstitutional), but that leaves the agency the power to compel behavior in the absence of a challenge, and even if arbitrary and capricious review is not as deferential in this context as it is with asset freezes (because of the national security dimension of those cases), it still leaves the court applying a fairly deferential standard of review to a decision that can be based on “any available information.” It’s quite plausible that the extent of resulting stringency in review is a reasonable compromise if the standard is going to be applied across the board, to every compliance order. It’s also quite possible that such review will not say much about the quality of compliance order decisions, which could (or perhaps should) ultimately affect the extent of confidence in the regulatory structures. Put differently, better review of regulatory decisions to impose compliance orders could change the bundle of substantive powers and penalties that could be acceptable to an enacting legislative and interest group coalition.

⁹² *See id.*

⁹³ Clean Air Act (“the Act”), 42 U.S.C. §§ 7401-7671q.

⁹⁴ *See* 42 U.S.C. § 7413(b) and (c)(2000). For a detailed discussion, see generally Jason D. Nichols, *Towards Reviving the Efficacy of Administrative Compliance Orders: Balancing Due Process Concerns and the Need for Enforcement Flexibility in Environmental Law*, 57 ADMIN. L. REV. 193 (2005).

⁹⁵ *See* 42 U.S.C. § 7413(b)(2).

⁹⁶ *See* *TVA v. Whitman*, 226 F.3d 1236 (11th Cir. 2003).

⁹⁷ *See* *Alaska Dep’t of Env’tl. Conservation v. EPA*, 53 U.S. 1186 (2003); *Alaska Dep’t of Env’tl. Conservation v. EPA*, 298 F.3d 814 (9th Cir. 2002).

F. Audits Can Ameliorate the Limitations of Traditional Judicial Review.

The preceding examples hint at a basic problem with managing executive discretion: unfortunately, sometimes judicial review is bound to fail when it comes to striking a balance between the aforementioned costs and benefits in the preceding domains. Figure 3 indicates the problem. What if a reasonable amount of review is simultaneously prohibitively expensive and extremely valuable? Imagine, for example, a situation where the cost-stringency function actually looks quite different from the probability-stringency function. Suppose, for example, that the first function is basically a linear one, just as in Figures 1 and 2. Stringency pushes up the cost of review in a linear fashion. But (as the lower portions of Figure 3 indicate) imagine that the relationship between stringency of review and the probability of discovering problems is starkly *discontinuous*. That is – a substantial increase in the stringency of review yields a very small appreciable benefit in terms of $Pr(\pi)$, even though the cost is still increasing in a linear fashion as the stringency does. Then suddenly there is a substantial increase. This sort of relationship might describe situations where most of the problems affecting targeted discretionary decisions involve complicated willful malfeasance rather than simple mistakes. The decisionmakers might engage in a concerted effort to cover up their tracks – which could only be detected with an extraordinarily thorough investigation.

[FIGURE 3 HERE]

Now put this together with a relatively low maximum acceptable cost (S_{max}), which is what the court may set if the decision has to do with national security. And notice what this does to the feasible review set. Essentially, it disappears. This implies that at least in some instances where targeted discretion is used, there may be *no feasible review set*, because the cost of reviewing discretionary decisions at the level of stringency necessary to actually have a credible chance of detecting problems far exceeds the maximum socially-acceptable cost of review.

Things look a little different if we use audits instead of traditional judicial review. The right side of Figure 3 tells the story. As with Figure 2, the slope of the cost-stringency line is flatter (more elastic), which leads to a higher stringency of review for any given cost that is paid for review. The result creates a small feasible review set – enough room for the auditor to avoid exceeding the maximum cost yet still have a quite reasonable opportunity to detect problems afflicting discretionary decisions.

When it comes to detecting those problems, audits hold another advantage over traditional judicial review. By its nature, litigation produces a biased sample. The cases we learn about are the ones that get litigated, and under various plausible conditions those cases are not the only ones likely to involve valid claims.⁹⁸ As Bill Simon's pointed out in the welfare context, for example, this pattern can seriously distort our understanding of a policy's strengths and weaknesses. It can also lead to court decisions may privilege

⁹⁸ There may be both overbreadth and underinclusiveness problems. For an early discussion of selection effects, see George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1 (1984). For an illuminating discussion of how different institutional rules can shape selection effects, see Peter H. Huang, *Lawsuit Abandonment Options in Possibly Frivolous Litigation Games*, 23 REV. LITIG. 47 (2004).

existing claimants over potential ones.⁹⁹ This doesn't mean that litigation and judicial review are not valuable in other ways. They can vindicate the interests of people who are legally and morally entitled to a proverbial "day," in court, or to a set of special remedies for which litigation is the best rationing device. Litigation harnesses the intricate machinery of adjudicatory bureaucracy to articulate and clarify legal norms in the context of specific cases. Rulemaking could let agencies "move from vague or absent statutory standards to reasonably definite standards, and then, as experience and understanding develop, to guiding principles, and finally, when the subject matter permits, to precise and detailed rules."¹⁰⁰ But these modes of inquiry are like discretion itself: replete with strengths and weaknesses. They should be treated accordingly and supplemented with audits.¹⁰¹

Notice that getting these benefits depends on making at least a few assumptions about the powers of the "auditor," though I suspect they are not entirely implausible. Specifically, the auditor needs to be in a better position to discover problems in the use of targeted discretion than the bureaucracy being reviewed. This probably means three things: (a) that the auditor is *motivated* to discover problems (and not to exaggerate them), thereby avoiding some of the willful malfeasance and politically-oriented self-dealing problems that bureaucracies have because of their political context; (b) that the auditor has sufficient abilities to evaluate the discretionary decision, perhaps in part through reference to some explicit or implicit standard of what is expected from such decisions; and (c) that the auditor is at least somewhat better than the decisionmakers being reviewed at avoiding some of the more subtle mistakes that afflict discretionary decisionmaking. Tempting as it may be to collapse these conditions into an "expertise" parameter, it's important to recognize that expertise (in addition to being a far more ambiguous term than is often recognized) is a dangerously seductive yet potentially quite dangerous two-edged sword: what makes some bureaucratic decisionmakers blind to the complexities of the problems they face is precisely their expertise in defining those problems in a standard, predictable fashion that often turns out to be wrong.¹⁰² Obviously the auditor(s) must know something of what it means to correctly freeze assets or assess regulatory penalties, but it's not obvious if that implies the sort of familiarity that allows a generalist judge to handle, say, criminal and complex tort cases, or whether it's more akin to the kind of specialization one already observes in the GAO and IG offices.¹⁰³ Whether it's possible to generate this and other conditions depends in large measure on how to resolve questions about the details of the institutional design, to which I turn below.

⁹⁹ See William H. Simon, *Rights and Redistribution in the Welfare System*, 38 STAN. L. REV. 1431 (1986).

¹⁰⁰ Culp Davis, *supra* note 1, at 219.

¹⁰¹ See *infra* Part II.b.

¹⁰² See, e.g., Vaughan, *supra* note 62, at 62 ("[T]he consequence of professional training and experience is itself a particularistic world-view comprising certain assumptions, expectations, and experiences that become integrated with the person's sense of the world."). See also CHARLES PERROW, *NORMAL ACCIDENTS* (1984); SCOTT D. SAGAN, *THE LIMITS OF SAFETY* (1995).

¹⁰³ And obviously, whatever the benefits of specialization (which help establish the presence of condition "b" above), they need to be balanced against the risk of sacrificing condition "c."

II.

THE INSTITUTIONAL DESIGN PROBLEMS ARE TRACTABLE

A. Audits Can Be Adapted to Address Multiple Forms of Discretion

The preceding examples and discussion suggest that a certain kind of discretion, which could be called “targeted discretion,” may be especially (though by no means uniquely) suited to audits. This is the kind of discretion showcased in the preceding examples. As Kenneth Culp Davis recognized a generation ago, the government is replete with power to make highly informal decisions affecting people, where “the usual quality of justice” may be quite low.¹⁰⁴ Such discretion, showcased in the preceding examples, primarily involves specific, targeted decisions whose primary effect is on specific individuals and groups.¹⁰⁵ Similar decisions involve bureaucracies applying some implicit or explicit legal standard, often in combination with some sort of policy basis (i.e., “enemy combatants are dangerous terrorists, many of them linked to Al-Qaeda”) that the executive branch itself has articulated as a rationale for these decisions.¹⁰⁶ Targeted discretion stands in contrast to broader policy judgments. Those judgments involve questions of how to interpret a statute or the relevant policy considerations when developing a legal standard, such as a regulatory rule or the content of an executive order.

Obviously, some policy judgments are designed precisely to affect targeted discretion. The Social Security Administration promulgates standards governing benefit payments, thereby making a policy judgment about *how* to use its targeted discretion. When government freezes allegedly terrorist assets, the State and Treasury Departments implement statutory standards from the International Emergency Economic Powers Act (IEEPA), as interpreted through policy judgments in the president’s executive orders. But despite their overlapping contours, these different sorts of discretion should nonetheless be distinguished. Though others may disagree, it seems as though policy judgments call for a different kind of review compared to targeted discretionary decisions that so often involve applying rules or standards to a particular set of facts. Whatever the arguments for deferring to the executive branch when an agency writes a rule or a president signs an executive order, those arguments seem at least somewhat weaker when the executive branch claims to be applying a standard to the facts. In the latter case the implicit claim is: “we may have to apply some judgment, but when we detain someone as an enemy combatant, freeze assets, or inspect an industrial plant, there’s no question about the purpose we are serving. We’re enforcing the law.”

No doubt sometimes government officials will argue that the details of a policy judgment – like precisely what behaviors make a charity liable to have its assets frozen –

¹⁰⁴ Culp Davis, *supra* note 1, at 216.

¹⁰⁵ Targeted discretion bears some resemblance to the concept of “informal adjudication” long discussed by administrative law scholars. For a brief discussion of the definitions associated with “informal adjudication” and some of the doctrinal problems implicit in this category of administrative action, see Ronald J. Krotoszynski, Jr., *Taming the Tail That Wags the Dog: Ex Post and Ex Ante Constraints on Informal Adjudication*, 56 ADMIN. L. REV. 1057 (2004).

¹⁰⁶ See Brief for Respondent, *Hamdi v. Rumsfeld*, *supra* note 7, at 12.

should not be made entirely public.¹⁰⁷ They might also argue that standards reflecting policy judgments should develop organically in response to experience (what would Diane Vaughan say to *that?*), instead of being fixed ahead of time. Even if one finds these positions attractive on the surface, it seems easier to think about them if we accept at least *some* distinction between government decisions that fix a standard that is supposed to apply across cases and those decisions that *apply* standards (or even quite general values) to specific cases. Though some may even insist that certain discretionary decisions involving national security (for example) are entirely immune from any standard,¹⁰⁸ this seems hard to reconcile with a simple but persistent imperative: that government decisions should not be arbitrary.

Audits of executive discretion may be viable even outside domains where their utility is most readily observable. As I note in Part IV, it may be possible to modify audits to shed light on applications of discretion drawn from a sparse set of decisions, or on policy discretion exercised in the course of rulemaking, for example, may nonetheless still be suitable for some kind of modified audit.¹⁰⁹ Another complication is epitomized by cases like *Webster*, where existing law fails to define a standard against which conduct can be assessed.¹¹⁰ Obviously, there would not be much discretion to audit if the domain were restricted to instances where a clear legal standard already existed. Ironically, in many circumstances where statutes are exceedingly vague – as with the Sherman Antitrust Act – a standard nonetheless emerges from the process judicial review itself. The absence that in some contexts is precisely what makes audits desirable. In the section that follows I discuss how such a standard might be derived by the auditor from overt executive—branch assertions or through reasonable interpretations of related constitutional, statutory, and policy principles.

B. Most Institutional Design Problems Have Plausible Solutions

Lurking just beneath the surface are practical questions of institutional design. My goal here is mostly to acknowledge those questions and begin thinking about how they might be solved. Resolving them completely would take some considerable attention to the particular kind of targeted discretionary decision that needs to be

¹⁰⁷ See *Doherty v. U.S. Dept. of Justice*, 775 F.2d 49 (2d Cir. 1985).

¹⁰⁸ The contention may be that, in addition to being free of any sort of formal review, certain national security decisions – like a presidential choice to bomb a potentially threatening target in the Sudan – *should not* conform to any standard at all. This position seems to confound the question of whether we should avoid setting a standard because it's too difficult to monitor it, or whether in an ideal world the president should never rely on a standard at all. The former is easier to justify than the latter.

¹⁰⁹ Small samples are not a fatal problem when Bayesian techniques are applied and the analysis is accompanied by appropriate assumptions. See, e.g., March, *supra* note 63; M. Elisabeth Paté-Cornell, *Organizational Aspects of Engineering System Safety: The Case of Offshore Platforms*, 250 SCIENCE 1210 (1990). Regarding policy discretion: although policy discretion arguably raises different problems, it may also be worth scrutinizing, case-by-case, the basis for application of an overtly stated policy; for example, how specific permissible exposure limits get set when courts allow rules that fix multiple standards – which they now discourage, though audits might be a promising alternative to simply banning such rules; instances of information dissemination and use regulatory agencies (as an alternative to the potentially cumbersome data quality act rules).

¹¹⁰ See *Webster*, 486 U.S. at 602-04.

reviewed, and to the political and institutional context affecting the agency (or agencies) making those decisions.

One question, for example, is how many decisions should be reviewed. The short answer is it depends on the total costs that can be incurred during the review process (both in terms of direct costs and in terms of the cost of limiting discretion). In keeping with the assumption that review increases as a function of the number of cases reviewed, the lower the “budget” for reviewing, the smaller the number of cases that can be reviewed. But if the budget is flexible, then how many cases would be “enough”? Think about this problem by envisioning two different kinds of situations that might arise. Sometimes there may be thousands, or tens of thousands, of targeted discretionary decisions to audit – as with border screening decisions or container inspections. In those situations, we might apply some of the counter-intuitive insights from probability theory. Those theories suggest that the accuracy of inferences drawn from analyzing a sample of cases depends on the number of cases sampled, and not necessarily on the total proportion of cases drawn from a population.¹¹¹ A sample of 500 cases, for example, might be associated with a confidence interval of plus or minus 6%, and one of a thousand cases drives the range down to about plus or minus 3%. The number of cases audited from a large number of cases would then depend at least partly on the desired confidence interval. But other factors would also matter. The more it is important to constrain perceived executive abuses, the greater the number of cases that should be audited. One might even envision a sliding scale, where the proportion of cases audited *grows* in response to the results of earlier audits with a smaller number of cases. All of these considerations then need to be weighed against the two kinds of costs (direct and indirect) that review implies.

What if we’re dealing with domains with smaller numbers of cases (say, a population of less than 1000 cases, as with Specially-Designated Terrorist Organizations)? Here, cost constraints could so limit the number of audited cases that it would not be realistic to expect very reliable confidence intervals. Nonetheless, audits would still serve the useful purpose of constraining executive authorities, who would know that some proportion (ranging, initially, between 10 and 20 percent depending on cost) of their decisions would be carefully scrutinized to make sure they’re done effectively.¹¹²

In both cases, one important question is whether the number of cases would include *both* instances where targeted discretion resulted in some sanction or cost being imposed (creating the potential for so-called Type I errors), as well as those instances where it was *not* imposed (Type II errors). For example, should someone audit just those cases where a charity *was* labeled a Specially Designated Terrorist Organizations, or also those where sanctions were *not* imposed? It’s quite likely that we would learn a good deal from including the cases where powers were *not* used. But this would raise two

¹¹¹ See A.I. Dale, *On a Problem in Conditional Probability*, 41 PHIL. OF SCIENCE 204 (1974).

¹¹² See *supra* note 110. Moreover, it may be possible to deal with domains involving extremely low numerosity by creating a reliable system to randomize the probability that a decision will be reviewed at all (or, to put it differently, by lumping together several classes of related, low-numerosity decisions and then auditing some proportion of them).

problems that need to be resolved later. First is deciding whether the expanded population of cases should include the whole universe (i.e., every charity, or – say – every charity operating in the Middle East) or just “near misses” (charities that attracted the attention of State, Treasury, the CIA, or the NSC but, perhaps because of political considerations, were not specially-designated). The former is more accurate but would so quickly consume auditing resources that it may not be feasible. The latter is simpler but less accurate. The second problem is overcoming the likely political resistance (from the executive branch, who would *already* have reason to resist audits) that comes from expanding the scope of review even further. This is a more general problem to which I return below.

Another recurring question is who exactly would do the audits. Some observers may insist that courts may lack the inclination, legal authority, culture, or expertise necessary to engage directly in audits, though they could probably appoint masters to do some of this work and they could fashion doctrines conditioning deference on the existence of reliable auditing done by someone else, or providing for audits as a remedy in the (unlikely) case where litigation itself reveals bureaucratic failures.¹¹³ By rewarding bureaucracies with reliable audit structures, courts could advance two interrelated objectives. They could contribute to mechanisms likely to enhance the overall quality of discretionary decisions (relative to some defensible, socially-relevant standard of quality encompassing, for example, reductions in the probability of obvious mistakes), and they would be creating the conditions for enriching the information on the basis of which a court can resolve specific cases. To the extent that courts run into *Vermont Yankee* problems,¹¹⁴ legislators could create an article I court with a distinctive mission and resources to build specialized capacity – or an entirely separate bureaucracy. Among existing agencies, the GAO and IG Offices are best positioned to do this sort of work (though, as I note below, they have largely avoided doing so). In short, a combination of judicial and legislative innovation could lower barriers preventing Article III courts from more easily encouraging audits, which could also be performed by existing federal audit bureaucracies. I discuss the political challenges to these scenarios in Part IV.

Whatever the institutional structure, the auditor (let’s use that term for now) would probably need the power to compel production of evidence and testimony, along with a cadre of independently-funded investigators. Otherwise it would be hard for the auditor to delve into enemy combatant designations or container inspections more aggressively than a court could. Sensitive information could be reviewed in-camera (something that would further weaken the argument that review is impossible because the information involved is too sensitive).¹¹⁵ Because this problem has been so often

¹¹³ They would have to do this in a way that avoids running afoul of the Supreme Court’s *Vermont Yankee* decision. See *Vermont Yankee*, 439 U.S. at 961.

¹¹⁴ *Id.*

¹¹⁵ We already let judges sometimes handle sensitive information and use it to make decisions in cases. See *U.S. v. Nixon*, 418 U.S. 683 (1974). It’s hard to see why audits of targeted discretion should be avoid on the grounds that the underlying discretionary decision depends on sensitive information. If the argument is that it’s dangerous or problematic to share the *results* of audits of targeted discretion because the policy domain requires complete secrecy even of the *quality* of decisions being made – then that argument should be advanced on its own merits and it should have to overcome a high barrier. See, e.g., *Ex Parte Quirin*, 317 U.S. 1 (1942).

managed in other contexts, I suspect any objection to audits relying on it is a red herring. Recent history is full of examples where this problem has been solved.¹¹⁶ In addition to courts reviewing the information in camera, high profile commissions like the 9/11 Commission and expert working groups routinely get security clearances and access to classified information.¹¹⁷ The resulting, publicly disclosed work product either omits classified information or provides some redacted summary version of it.

Regardless of whether the case involves sensitive information or not, what *standard* would the auditor use to evaluate it? Ideally the statutes or constitutional provisions implicated in the discretionary decisions would provide *some* standard for the auditor to use, even when the standard is too vague for courts to apply. Or the auditing authority can analyze whether a number of statutes and constitutional doctrines together could be taken to imply conditions on the use of discretionary powers.¹¹⁸ The auditor could even use statements from the executive branch itself to see whether the audited cases seem to be consistent with those statements.¹¹⁹ In some circumstances, where the executive refuses to articulate an explicit standard to fill in gaps left by executive, legislative, or judicial silence, the auditor *itself* could articulate a reasonable standard (which is, by the way, what the GAO and IG do in related context, when they audit “broad management practices”). The standard might reflect insights drawn from constitutional interpretation, policy considerations, or even statutes’ legislative history.

A final issue is what to do with the results of the audits. The question turns on whether audits are expected to provide an independent remedy for aggrieved individuals and groups – like a court proceeding would – or whether they are simply expected to inform the judgments of legislators and the larger public. One possibility is to announce the results of the audits (minus the sensitive information). What this would accomplish

¹¹⁶ See, e.g., *Jifry v. FAA*, 370 F.3d 1174 (DC Cir. 2004).

¹¹⁷ See REPORT OF THE NATIONAL COMMISSION ON TERRORIST ATTACKS UPON THE UNITED STATES (2004).

¹¹⁸ Deriving such “implicit” standards is exactly what the Supreme Court has discouraged in cases like *Heckler v. Chaney*, 470 U.S. 821 (1985). There the Court held that agency decisions not to use their enforcement powers are almost always “committed to agency discretion” under the terms of the Administrative Procedure Act, 5 U.S.C.A. §§ 701-706 (2000). The exception is where a statute provides “clear guidelines” that a court could use as a standard against which to judge agency decisions. Indeed, when no such standard is apparent on the face of the statute, then courts tend to find that the absence of such a standard overcomes what is otherwise a presumption of reviewability. See *Lincoln v. Vigil*, 508 U.S. 182 (1993). What these cases sometimes gloss over (but seem to recognize far more explicitly in cases involving the nondelegation doctrine) is the extent to which the existence of standards is on a continuum, where virtually any government action or inaction (including decisions not to prosecute) could be evaluated in accordance with some defensible criterion. For an exception, see *Adams v. Richardson*, 480 F.2d 1159 (DC Cir. 1973)(agency has “consciously and expressly adopted a general policy” that is so extreme as to amount to an abdication of its statutory responsibility). This is not to suggest that courts should be the primary actors in conducting targeted discretion audits (perhaps some independent auditing authority including a mixture of judges and non-judges would have more flexibility to articulate standards). It is, rather, to point out that courts’ reluctance to articulate standards when they find them missing on the face of a statute should not be taken as an indication that such an enterprise is fruitless.

¹¹⁹ For example, if statutes say that the CIA Director can fire someone for being a national security risk and let him define what that means, then the audit could review the *definition* with particular care to see if it’s plausible – and announce the results. Or it could rely on agency definitions of “national security” in related contexts. Cf. *Chaney*, 470 U.S. at 836 (suggesting that an agency decision not to enforce could be challenged if the agency itself has committed to act in specified circumstances).

depends on the reactions of legislators and the mass public, which can vary depending on the circumstances.¹²⁰ Another possibility is to give relief (or impose belated sanctions) whenever audits reveal problematic cases. A woman improperly barred from entering the country could be allowed to return. Assets that should not have been frozen could be unblocked. Enemy combatants could be set free. This is certainly a principled position, though it obviously raises certain costs associated with the audits.¹²¹ A third approach is for the results of audits to trigger additional procedural standards, such as review of more decisions and (perhaps) greater judicial scrutiny. The choice among these alternatives is likely to depend heavily on some of the political considerations I discuss in Part III. What's important is to think of the choice not only in terms of what benefits could be provided to aggrieved individuals, but also (more generally) how different remedial schemes are likely to impact agencies' willingness to learn from their mistakes and structure its work to avoid future abuses. In short, through careful institutional engineering, analogies to existing institutions, and some experimentation, most of the "problems" identified can be solved. We might then ask whether they have been solved already.

III.

REFORMS CONTRAST SHARPLY WITH THE STATUS QUO

The value of supplementing judicial review with audits of executive discretion depends largely on whether or not something like these audits already occurs. Audits of targeted discretion may sound like exactly the sort of work that the GAO and the IG Offices already do. These audit bureaucracies were, after all, created to audit the government, and their jurisdiction has expanded to include investigating the management of government programs.¹²² Their activities are sometimes directed by legislators, who (in turn) can proceed with their own audits. Do they?

A. Federal Bureaucracies Do Not Ordinarily Perform Audits

For the audit bureaucracies the answer is basically "no." These agencies do a tremendous amount of interesting (and probably sometimes quite valuable work) on bureaucratic performance. The scope of their authority is quite broad. The GAO, for example, has the power to examine "any matter" relating in some way to the disbursement of public money.¹²³ The Inspector General offices in federal departments have a similarly broad mandate.¹²⁴ Yet they rarely seem to do rigorous audits of targeted executive discretion.

¹²⁰ See Asher Arian, *Political and Economic Interactions with National Security Opinion: The Gulf War Period in Israel*, 43 J. CONFLICT RES. 58 (1999); David O. Sears et al., *Political System Support and Public Response to the Energy Crisis*, 22 AM. J. POL. SCI. 56 (1978).

¹²¹ Although I do not develop the argument here, one might imagine a situation where the designers of the audit system would trade-off the ability to grant relief in exchange for the political and economic resources to audit more cases or to do so more intensely. That's the sort of bargain that supporters of "truth commissions" are willing to make.

¹²² See *infra* notes 124, 125, and 142.

¹²³ 31 U.S.C. § 712 (1).

¹²⁴ See Inspector General Act of 1978, Pub. L. No. 95-452, 92 Stat. 1101 (codified as amended in scattered sections of 5 U.S.C.).

There is some literature on the historical origin, legal jurisdiction, organization, and culture of these audit bureaucracies. But we know relatively little about what the reports of these audit bureaucracies are about, what methods they use to develop their analyses, whether these reports contain recommendations that agencies actually implement, and whether any of this gets media attention. These questions are obviously relevant here because they affect whether there is a deficit of the kinds of audits I recommend, and their answers help us learn something about how audit bureaucracies might go about doing more of what I recommend. To begin engaging some of these questions, I can present some initial results from a more extensive ongoing study of the audit bureaucracies. The data were obtained primarily from the GAO and IG Offices websites, or (in some cases) directly from the GAO. Four trained research assistants have worked with me to code many of these reports and gather additional data about key agencies. Among other things, our goal has been to see how frequently the GAO and IG Offices do targeted discretion audits of the kind I recommend above. The results, shown on Figure 4, indicate how rare this is. The figure divides reports on the basis of whether they do something like a targeted discretion audit. Only a small number of reports appear to do anything even remotely resembling an audit of executive discretion, and even among those reports the focus is only on a small subset of the discretionary authority. Some reports occasionally chronicle problems in administrative systems like those governing aviation security. But audits of discretion are altogether missing for a number of categories of decisions, including “no fly” lists, enemy combatant designations, individuals removed at the border and not allowed to apply for asylum, and administrative compliance orders from agencies like OSHA. For instance, the GAO’s otherwise thorough report on the Transportation Security Agency’s computerized aviation security system shows how carefully the agency reviewed the architecture of the computer algorithm and the management practices associated with the systems. It did not, however, pick a subset of names to inquire exhaustively how they ended up on the list or what evidence supported that determination.¹²⁵

[FIGURE 4 HERE]

My preliminary data on the GAO and IG Offices also lets me anticipate and offer a partial response to two skeptical views about the idea of audits. First, would anyone in the public actually *care* about the results of the audits? After all, much of the idea is predicated on the notion that there may be political consequences to revealing how discretion is being used. This is particularly important if the scheme is not designed to deliver a direct remedy to people affected by the reviewed cases. Second, wouldn’t agencies simply ignore the results of the audits? If they did, then the promised learning bonus from audits would be unlikely to materialize.

Even at this early point in the empirical phase of the project, it appears as though GAO and IG Office reports get a considerable amount of attention in the print and television media. The following figure shows the number of stories in the New York Times and in transcripts of television news stories between January 2002 and January 2005 mentioning the GAO or IG Offices. Figure 5a shows the number of New York Times articles mentioning the audit bureaucracies. Nearly a thousand do. A random

¹²⁵ See GAO Report, *supra* note 83.

sample of 200 of those news stories indicates that, while only about 3% of the stories involving the GAO appear on page 1, about 10% of those mentioning the IG Offices do so. Audit bureaucracies are also discussed on broadcast news and cable channels. Even in these media, nearly a hundred news segments mention the various IG Offices, and about 30 mention the GAO.¹²⁶

[FIGURES 5]

What about recommendations from auditors being ignored? I have obtained data on *all* the 10,000 or so recommendations made by the GAO over the last 15 years. After the GAO makes recommendations to an agency in its reports, what it does is to follow up through interviews, additional investigation, document reviews, and queries to the agency leadership. It then determines (on the basis of these qualitative methods) whether a given recommendation is implemented sometime during the next four years.¹²⁷ It turns out that a whopping 79% of recommendations are implemented. Obviously there's a lot more work to be done on these recommendations. For example, the extent to which recommendations are adopted may be endogenous to *what* the recommendation is – with simpler ones (i.e., “write a report on the quality of the vehicle fleet for the Secret Service”) being implemented much more than complicated or difficult ones (“reduce the extent to which the Secret Service works on simple credit card fraud cases instead of critical infrastructure protection”). I strongly suspect that the adoption of recommendations is influenced by political factors, such as the extent of division in appropriations and authorizing subcommittees that oversee the agency in question. It's also quite likely that departments with different bureaucratic structures, institutional cultures, and particularly those with greater prestige, have different reactions to the GAO recommendations. What makes little sense is to blithely reject the relevance of these audit bureaucracies – even if they don't currently do the sorts of audits that would help us learn more about the use and abuse of targeted discretion.

B. Neither Do Legislators

Another possible setting where targeted discretion audits could take place is in the legislature, where hearings to oversee the bureaucracy are routine and legislators often complain loudly about what agencies have done. As it turns out, most legislative oversight activity has virtually nothing to do with systematically auditing targeted discretion. In Part III I suggest some of the reasons why, as with the audit bureaucracies, there seem to be so few audits of targeted discretion. In what follows I just want to provide a brief outline of what legislative oversight activity tends to look like, and how this is different from targeted discretion audits.

Soon after legislators arrive in Washington, many of them almost invariably find that “oversight” of the bureaucracy has big payoffs. It lets them achieve desired policy goals. It also lets them claim credit for making the government work more efficiently and effectively. As a consequence, legislative oversight activity takes on a bewildering

¹²⁶ The search involved review the transcripts of all news stories on CNN, CNBC, CBS, NBC, ABC, and Fox News available on Westlaw, with airdates between January 2002 and January 2005. Transcripts of local news programs would probably pick up additional stories.

¹²⁷ Interview with GAO Official # 1 (January 2005). The agency (quite plausibly) assumes that if a recommendation is not implemented within four years of being made, it probably will not be.

array of forms, including – among others – formal committee and subcommittee hearings, staff investigations of bureaucratic practices, direct contact between a legislator and an agency’s leadership, meetings with the White House to enlist its support in pressing a bureaucracy into service, and control of the appropriations process.¹²⁸ In the mid-1980s, political scientists Mat McCubbins and Tom Schwartz introduced what has become an incredibly durable framework for thinking about legislative oversight of the bureaucracy.¹²⁹ Police patrol oversight involves legislators using their time, staff, and other resources to engage in fairly constant vigilance of agency outputs – primarily through staff investigations and committee hearings. In contrast, “fire alarm” oversight requires less constant attention from legislators and their staff. Instead, legislators wait for “fire alarms” to be pulled by interest groups and portions of the public (occasionally, perhaps when galvanized by media attention to some perceived regulatory problem). To encourage this sort of activity, legislators create procedures such as the federal Administrative Procedure Act and the Freedom of Information Act that let groups more easily learn what’s going on. Legislators rely on these parties to assist (implicitly) in the oversight process. In short, fire alarms involve two related features: (a) reliance on interest groups (or, on occasion, a politically engaged citizenry), and (b) episodic legislative responses to instances where these groups

The fire alarm concept is almost the opposite of a random audit. Unless legislators directly create a procedure to audit targeted discretion (they haven’t so far), then fire alarms would virtually never involve auditing, but rather sharp responses when problems have already surfaced. Moreover, because targeted discretion often (though not always) affects individuals or groups without ready access to political power, fire alarm oversight would be particularly unlikely to uncover problems. In contrast, police patrol methods are much more consistent with the kind of audits I describe. Yet there is little evidence from congressional testimony and hearings that this is the sort of oversight that legislators do directly. In fact, what their public statements seem to suggest is that if anyone is doing the kinds of audits that reveal problems with government, it’s the GAO and the IG Offices, not their own staff.¹³⁰

No doubt that congressional investigations often uncover important trends or problems in bureaucratic activity, whether they are triggered by fire alarms or they arise from more pervasive police patrol methods. Nonetheless, those very same investigations also run into trouble getting access to information about how discretionary decisions are made. Moreover, while it is perhaps it’s true that some forms of legislative control can substitute for other mechanisms – like audits – two basic facts might nonetheless make audits of executive discretion distinctive compared to most of what legislatures, courts, and audit bureaucracies currently undertake. (1) Legislators train their attention on what catches their attention, not on a random sample of discretionary decisions. Decisions that are not reviewed randomly provide a biased sample, which can skew the reported results either because of inherent characteristics in the sample or because the players being “audited” strategically distort what they’re doing in the decisions more likely to be

¹²⁸ See JAMES Q. WILSON, *BUREAUCRACY* (1990).

¹²⁹ See Mathew McCubbins and Thomas Schwartz, *Congressional Oversight Overlooked: Police Patrols vs. Fire Alarms*, 28 AM. J. POL. SCI. 165, 166-68 (1984).

¹³⁰ See, e.g., Light, *supra* note 42, at 42

audited. Cases that are not reviewed at all don't become the subject of any legislative, political, or public pressure. We don't learn anything if we don't review *something* relevant to the discretionary decisions made constantly throughout the government. (2) Even when legislators and their staff choose to focus on something, their oversight does not necessarily imply review of specific decisions. As with the audit bureaucracies, oversight hearings may focus on systemic issues such as an agency's policy priorities or its handling of obvious crises. While staff may occasionally review random samples of case files, this is not a routine component of legislative hearings. From a prescriptive point of view, the results may provide less explicit – and instead more ambiguous – findings, which are harder to interpret and have less to say about whether government is performing effectively.

IV.

AUDITS FACE POLITICAL CONSTRAINTS

Judicial review is likely to do a poor job of overseeing targeted executive discretion, because the standard of deference is often set too high to give a meaningful chance of discovering problems. Audits could tell us more, at a reasonable cost. But they rarely happen. Why?

A. Observers Sometimes Fail to Distinguish Between Targeted Discretion and Policy Judgments

I began this article by distinguishing among different kinds of discretion. Some of government's work is about writing executive orders to determine when covert operations are warranted, designing regulatory rules for disposing of nuclear waste, or similar tasks that involve broad policy judgments. Other decisions involve discrete choices to impose costs (or deliver benefits) primarily on specific individuals and groups, generally in accordance with some implicit standard or broad policy goal. Most of my argument is about how to improve our ability to learn how bureaucratic institutions use the *second* kind of flexibility, targeted discretion, to shape our welfare and security. Although I have already acknowledged and discussed some overlap between such targeted discretion and policy judgments, they are really quite different. Policy judgments are about interpreting statutes, ambiguous facts, or policy considerations that apply across an entire class of cases. Targeted discretion involves a discrete choice that allocates burdens or benefits primarily onto an individual or group. Judgments of the former kind lie at the core of regulatory rulemaking, on the other hand, are not. The latter kind of discretion is instead primarily about applying legal rules, implicit standards, or policy judgments onto specific cases. Indeed, even when targeted decisions seem entangled with broad policy judgments, we can draw distinctions and make the former kind of decision subject to a different sort of review (one that might even take as its standard the policy judgment itself). Because courts and some commentators fail to fully distinguish between these two kinds of decisions, it's often unclear just why audits would be appropriate. This also blurs the distinction between formal audits and the many audit-like political mechanisms that review the actions of government bureaucracies, like

Inspector General reports on broad program characteristics and objectives, as well as congressional fire-alarm type investigations.¹³¹

Lumping together these two kinds of discretion has the unfortunate effect of making it seem harder to design an effective means of reviewing executive discretion. From a court's perspective, deference probably seems more attractive if discretion seems to be about policy, because courts are so often assumed to lack the accountability and expertise to make broad policy judgments across large numbers of cases. Earlier I suggested that it might be possible to audit policy judgments such as those underlying regulatory rules, but doing so would almost certainly require different procedures. This has not stopped lawyers, judges, and scholars from often discussing discretion as if it were *one thing*. And courts make deference decisions that draw no distinction between supporting cases cited that concern targeted discretion, and those concerning broader policy judgments.

B. Political Actors Have Polarizing Incentives

A second likely reason why audits of targeted discretion have so rarely materialized thus far has to do with the incentives shaping the behavior of two important sets of players who have stakes in the work of the executive branch. Officials in the executive branch (and their allies in the legislature) obviously matter because they could institute an audit system internally. They could support its implementation by the GAO and IG offices. Or they could advocate for it in the legislature. The other set of players involves those legislators (and their allies among organized interest groups) who are generally opposed to expansive power in the executive branch.¹³² The problem, in a nutshell, is this: Executive authorities and their allies will tend to want more discretion, and less review. Unless they have a good reason for trying to cut down on their flexibility in order to send a costly signal of some kind to overcome a commitment problem,¹³³ these players would probably prefer to avoid the embarrassment of an audit that does not show them succeeding, and to retain the benefits of the flexibility implied by that discretion. On the other hand, those who advocate *against* executive power may gain something from keeping the debate polarized, which allows them to retain an issue that can galvanize supporters. And depending on the situation, some such advocates might prefer to hold out for winning more stringent review *across the board* instead of settling for a compromise.

Executive branch officials would be loath to part with discretion, for at least two reasons. First, discretion helps executive branch authorities carry out the functions that they are expected to, like keeping threatening people out of the country, prevailing in

¹³¹ See *supra* Part II (discussing the work of the GAO and IG Offices).

¹³² Obviously there are still other players who matter. There may be moderate legislators, for example, who have an interest in restraining abuses but fundamentally accept the argument that intrusive judicial review may be too costly. I suspect both that these actors could play a crucial and constructive role in promoting bureaucratic changes that could contribute to improvements in how we oversee targeted discretion. I also imagine that their numbers have been thinned by the increasing polarization of politics described by some recent work among political scientists. See Nolan McCarty, Keith T. Poole, and Howard Rosenthal, *Political Polarization and Income Inequality*, Princeton University Department of Politics: Working Paper (January 27, 2003)(on file with author).

¹³³ For a discussion of this possibility, see Cuéllar, *supra* note 65.

military operations abroad, or (at least some of the time) keeping industrial workplaces safe. Discretion is also valuable because it helps create certain impressions among superiors, legislators on appropriations committees, interest groups, and the public. People respond to what they can *see*. Targeted discretion lets government officials (or their subordinates) *choose* what seems to be happening in a given area of the law. It stands to reason that losing some of this power is not a welcome prospect. Neither is it desirable to face the additional costs and the possibility of embarrassment that come with more stringent audits. One might expect supporters of executive power in the legislature to take a similar position.¹³⁴

What about legislators and interest groups concerned about limiting executive power? Think first of the legislators who tend to distrust what the executive branch is doing. Political scientists have had a great deal to say about how legislators oversee the bureaucracy. A now-classic strand of literature distinguishes between “fire-alarm” oversight prompted by some interest group complaint or dramatic visible failure in the executive branch, and “police patrol” methods that involve steady monitoring of the bureaucracy week after week.¹³⁵ Police patrol methods seem to encompass audits. But as the literature persuasively establishes, legislators don’t always want to use that tactic because it is costly. *Even* assuming that the conditions are present to make these legislators want to use police patrol methods instead of just waiting for an interest group to complain, it’s not obvious that the critics of executive power would want to press for rigorous audits instead of simply polarizing the debate or attempting to embarrass their political opponents.¹³⁶ A highly polarized debate has some benefits. It may galvanize support among certain constituencies. And legislators (along with their allies in external interest groups) may prefer to win across the board than to support solutions that no doubt seem to some like flimsy half-measures.

One can tell much the same story about advocacy organizations outside government. If the issue is the treatment of enemy combatants, for example, Human Rights Watch may strongly prefer a system where authorities implement the *Hamdi* decision in a way that drastically cuts down on executive discretion. Audits may seem like a poor alternative by comparison. The choice between promoting audits (as a compromise) or pressing for a more stringent standard of deference across the board thus depends, as before, at least in part on the players’ subjective probabilities.

Considerably more could be said about the consequences of polarization for issues of legal and bureaucratic structure. The main point, for now, is that polarization probably diminishes the extent of political interest in review mechanisms that may be socially optimal. When players have more polarized views about executive branch power, substantive policy and law, or both, they probably have less to gain from investing in audits or similar approaches.

¹³⁴ Unless, of course their support of executive branch power is overwhelmed by incentives to support (for example) legislative power.

¹³⁵ See McCubbins and Schwartz, *supra* note 130, at 166-67.

¹³⁶ Of course, audits may in fact embarrass the executive branch. That depends on what the opponents think the audits will reveal. But whatever political benefits audits can provide must be weighed against the opportunity costs that I discuss above.

C. *Institutional Inertia Locks In Existing Conceptions of Adjudication*

Lawyers constantly think of themselves as zealous advocates on behalf of individual clients. Practicing lawyers join scholarly commentators in thinking of courts as direct protectors of the rights of similarly-situated individuals. These frames are in some tension with the idea of reviewing fewer cases in exchange for reviewing them more thoroughly. Adjudication is often perceived as embodying at least two important characteristics. First, if adjudication is going to be available, it is assumed to be a recourse that should be available to all similarly-situated parties. People who have been allegedly treated the same should not have different opportunities to vindicate their claims. Circuit splits are a fact of life, but academic lawyers consider them problematic precisely because they create distinctions in how similarly situated litigants would be treated. Second, adjudication is supposed to provide a remedy. Which is why judgments that don't provide a remedy strike some observers as ridiculous, and why some scholars have persuasively shown how it makes little sense to think about adjudication constitutional rights without "equilibrating" that adjudicatory process with the remedies in question.¹³⁷

As I've defined them, audits don't really conform to these assumptions. In a narrow sense, they randomly privilege some people and not others. They don't provide an obvious remedy (though it's certainly possible to fashion one). They seem, as a result, to be ill-fitting proxies for a persistent set of concerns that help justify less deferential adjudication. No doubt it's true that constitutional provisions and values may require adjudication. Sometimes deficiencies in adjudication are best remedied through changes in adjudication. But it seems equally clear that the prevailing conception of adjudication could unduly dampen interest in audits. It promotes the false sense that the value of audits are primarily seen where an individual abuse (or mistake) is discovered, and corrected. Instead, the point of audits is to shed light on the entire system and how it works. This has always been a concern of adjudication as well, but perhaps it sometimes gets lost amidst the pressing rhetoric about protecting individual rights. Courts inclined to serve as a counterweight can do so by crediting, during arbitrary and capricious or substantial evidence review, agencies who incorporate credible audits of their decisions, or who have been subject to such audits from the GAO and Inspector General offices recently. Although *Vermont Yankee* and similar cases preclude the full range of judicial elaboration of new procedures, it does not strain the existing scope of review to suggest that courts attend to the internal and external procedures shaping the extent to which a specific agency decision becomes arbitrary.¹³⁸

D. *Narrow Conceptions of Auditor Mission Prevail*

Which brings us back to how the audit bureaucracies behave. Like all bureaucracies, the GAO and IG Offices are also affected by prevailing conceptions of what it means for them to do carry out their work. Government employees who have some flexibility to choose what to do and how to do it tend to make choices reflecting – at least in part – their own sense of the mission they are supposed to carry out. Those

¹³⁷ See Daryl J. Levinson, *Rights Essentialism and Remedial Equilibration*, 99 COLUM. L. REV. 857 (1999).

¹³⁸ Cf. Lisa Schultz Bressman, *Beyond Accountability: Arbitrariness and Legitimacy in the Administrative State*, 78 N.Y.U. L. REV. 461 (2003).

choices can reinforce external perceptions, which in turn affect the work referred to the agency, the financial resources it receives, the people who apply for jobs there, and the standards used to evaluate whether the agency is succeeding in its work. Together these factors then combine with the more prosaic political pressures both within and outside the agency to shape its work environment.

The GAO and IG offices were created to serve as *auditors*. So at least one factor shaping the priorities of these bureaucracies would be how the role of an auditor is defined. As far as the legislators who created these bureaucracies are concerned, the reference to auditing seems to be taken primarily as a reference to detecting financial mismanagement or malfeasance. In the late 1970s, a GAO report commented on legislative plans to create IG offices, and emphasized the urgent importance of auditing the finances of government agencies. A scholarly commentator notes how this report emphasized the tenor of the congressional discussion at the time:

Surveying every unit of the federal government, from whole agencies to small program offices, GAO found that almost a third had not had a financial audit since 1974. In unusually dramatic prose on the front cover, the report announced: “One undred and thirty-three unites, with annual funding in excess of \$20 billion, told GAO they had not received a financial audit during fiscal years 1974 through 1976.”¹³⁹

The focus on financial waste and related themes should hardly be surprising, given the potential political payoffs of creating bureaucracies that are designed to get rid of “waste, fraud, and abuse.”¹⁴⁰ It’s not easy to find legislators or executive branch officials in favor of waste, fraud, or (financial) abuse – though (particularly for the GAO) it’s certainly plausible to think that the content and aggressiveness of investigations targeting such problems would change depending on the partisan composition of the legislature and executive branch.¹⁴¹ It’s also true that legislators may find audit bureaucracies useful to generate publicity and promote certain policy objectives.¹⁴² Nonetheless, to the extent that legislators expect these bureaucracies to audit, it seems that they primarily expect those audits to focus primarily on money and resources: who spent what funds, why government vehicles were used for that trip, or why these employees were asked to work on some questionable task. This is despite the wide jurisdictional latitude that the law gives both sets of agencies. There is no doubt that such latitude would allow them to effectively audit the quality of the targeted discretionary decisions, but doing so seems somewhat in tension with what legislators think of when someone says “GAO” or “IG.”

The same may be largely true of many managers and officials within the agencies themselves. Those officials obviously retain *some* discretion to make choices about what to audit (and what methodology to use) after all the external politics are taken into

¹³⁹ See Light, *supra* note 42, at 42.

¹⁴⁰ *Id* at 43-45.

¹⁴¹ See, e.g., Anne Joseph, *The Politics of Auditing*, Unpublished Ph.D. Dissertation, Political Economy and Government Program, Harvard University (2001).

¹⁴² See *id.* at 44 (noting how “some member[s] of Congress... use IG input to frame issues for resolution. However, some members of Congress have been less concerned with policy reform, using the IGs instead to score short-term political successes.”).

account. But too, it seems plausible that these agencies were locked in to a financial auditing mission. Many (though certainly not all) Inspectors General have a background in financial management or accounting. So does the Comptroller General, who heads the GAO. Combine this with the fact that both supporters and detractors of greater executive power have little incentive to push for audits of targeted discretion, and it starts to look like the audit bureaucracies would only focus on targeted discretion sporadically and idiosyncratically. And that's exactly what they seem to do.

CONCLUSION

Executive branch lawyers recursively emphasize how dangerous and complicated the world is when they argue before a court or submit a legal brief. So do the lawmakers who sell the President's program in the legislature. Both urge their audience to accept danger and complexity as the hallmarks of a world demanding more discretion and less doctrine. Harder to find among their verbal onslaught is any reference to two indelible facts. The first is that executive discretion always carries risks as well as rewards. There is nary a chance for sound governance, or any kind of governance, without discretion. Its existence lets government protect the environment, prosecute serial rapists, keep workers safe at industrial sites, and fight battles to protect its citizens. But history writes a damning indictment of discretion's abuse. It describes not only how Nixon's IRS embarrassed his enemies, or how Hoover's FBI libelously fed speculation that slain civil rights workers were promiscuous, mentally-ill subversives, but also how even the most determined and virtuous government officials fail to learn from their mistakes when they don't know they have committed them. None of this should be surprising given what we know about organizations, the people who run them, and the complicated legal mandates entrusted to them. The second fact is how much discretion the executive and its millions of workers *already* have.

This article considered the implications of these two facts in light of two other realities. Judicial review predictably and systematically fails to constrain the kind of executive discretion that most directly affects people and organizations (what I call targeted discretion). And we do next to nothing to audit how that discretion is used, despite the presence of at least three different compelling reasons to think that executive branch officials will have a relentless tendency to frequently misuse that discretion. (1) Because some discretionary actions can signal competence and resolve to naïve observers among the mass public, executive officials may have an incentive use their discretion to create favorable impressions. (2) Some officials or their employees may be far less subtle and engage in willful misconduct that is unlikely to be detected. (3) And organizations may have a harder time learning without external mechanisms to systematically review and critique their work.

It would be a mistake to reject the impact of traditional judicial review on many forms of executive discretion. But it would be equally misguided to assume that courts succeed in striking a reasonable balance between targeted discretion's benefits and its many costs. Instead courts make limited forays into the territory of executive power when they hear cases about targeted discretion. The more costly the courts or legislature assume those forays to be, the more deferential the review of targeted discretion. Whatever the merits of this approach, it creates a situation where existing methods of

court review can serve up the worst of both worlds: no meaningful probability of uncovering problems in discretionary decisions, coupled with a prevailing sense that courts can still remain the nation's conscience and protect us from discretion's darker side.

Many who have raised these concerns have done so by taking part in a familiar debate about the value of greater judicial scrutiny of executive discretion. While this article does not dismiss the value of that greater scrutiny, it offers an alternative to the polarized rhetoric of that debate. It effectively says: *even* if one accepts that more stringent judicial review is impossible, one should *not* therefore accept that the correct result is to let the executive branch's wheels keep on spinning as they always have. The key to that alternative is to think about the problem of policing executive discretion as an *information* problem. The value of this approach lies in its potential to help sever the connection between the perceived costs of encroaching on discretion (both in terms of direct review costs and in terms of interference with the valuable characteristics of discretion) and the stringency of review. This property makes audits a promising conceptual alternative to judicial review in those instances where it would be preferable to study a subset of cases more carefully instead of superficially reviewing all of them. Government powers to inspect, fine, prosecute, enforce, and detain may rightly seem less threatening if their use can be effectively monitored through audits or similar procedures.

It may seem at first as though audits would only work if we lived in a world perfect enough to make them unnecessary in the first place. But the institutional design problems associated with auditing executive discretion call for an altogether subtler diagnosis. Instead, four dynamics help explain that continuing embrace of judicial review, and the concomitant absence of activity auditing targeted discretion. When lawyers and policymakers erase the distinction between targeted discretion and broader policy judgments, they unduly restrict the scope of options available to help balance discretion's benefits and costs. Which is just fine for presidential administrations, executive officials, and legislators supporting executive power: They're perfectly happy to let that power evade more frequent review. Somewhat counter-intuitively, advocates of restraining that power may also have incentives to oppose audit-like approaches as a matter of political strategy, because it lets them sound the alarm to their supporters while they fight for more aggressive review across the board. That fight happens in a context permeated by persistent (yet ultimately misleading) norms about the appropriate relationship between adjudication and review of executive discretion, and similarly durable conceptions of what existing auditors should do when they supervise government agencies. Weakening these dynamics may require propitious circumstances and Herculean feats of advocacy, but not the "perfect world" that would let us dispense with audits (or, indeed, judicial review) altogether.

All of this tells a larger story that is not just about audits, but about the forces affecting what we consider to be plausible conceptual alternatives to judicial review when we address discretion-related problems. Judicial review gives judges, lawyers, and legislators considerable room to innovate and experiment with different kinds of review mechanisms. On the other hand, traditional court review does some things better than others, and in some domains it does little or nothing at all. Yet we cling to it as the premier method, other than perhaps the ill-defined political process itself, for providing

some kind of constraint on executive discretion. We can easily spend another ten months, or ten years, with the palliative machinery of the status quo grinding on. So what? Surely it is naïve to assume reflexively that each unchecked discretionary decision amounts to a disaster. What borders on madness is to think those decisions will turn out just fine when existing law lets them so easily escape scrutiny. Perhaps it will take decades to know the full costs of the inevitable mistakes and manipulation that get swept under the rug when we freeze the wrong charity's assets or inspect the wrong containers. Then again, we may not be so lucky.

Figure 1: Discretion and Traditional Judicial Review

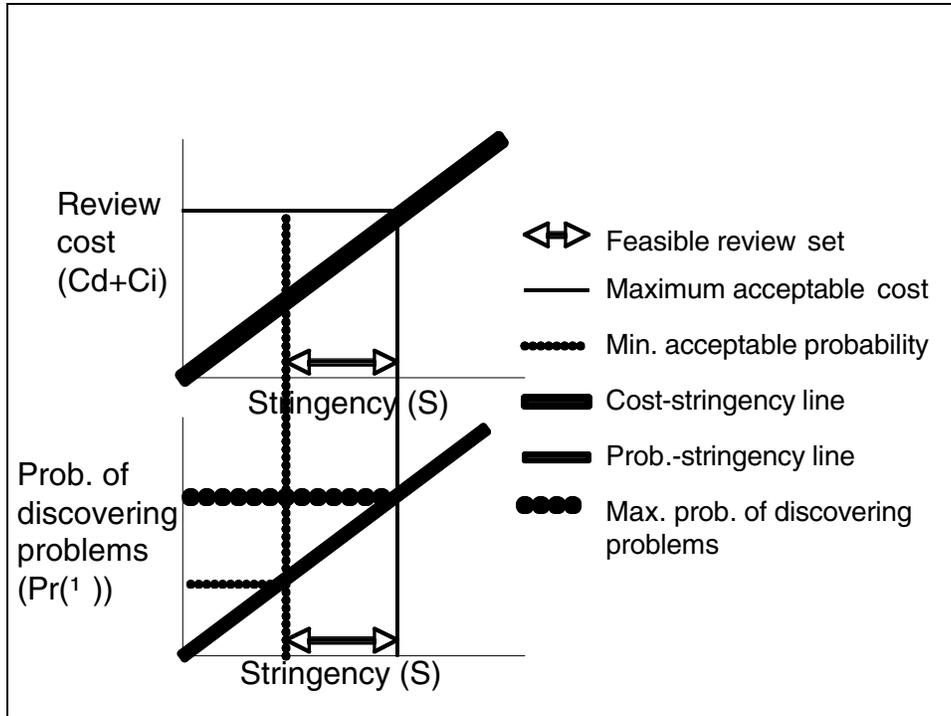


Figure 2: Comparing Audits to Traditional Judicial Review

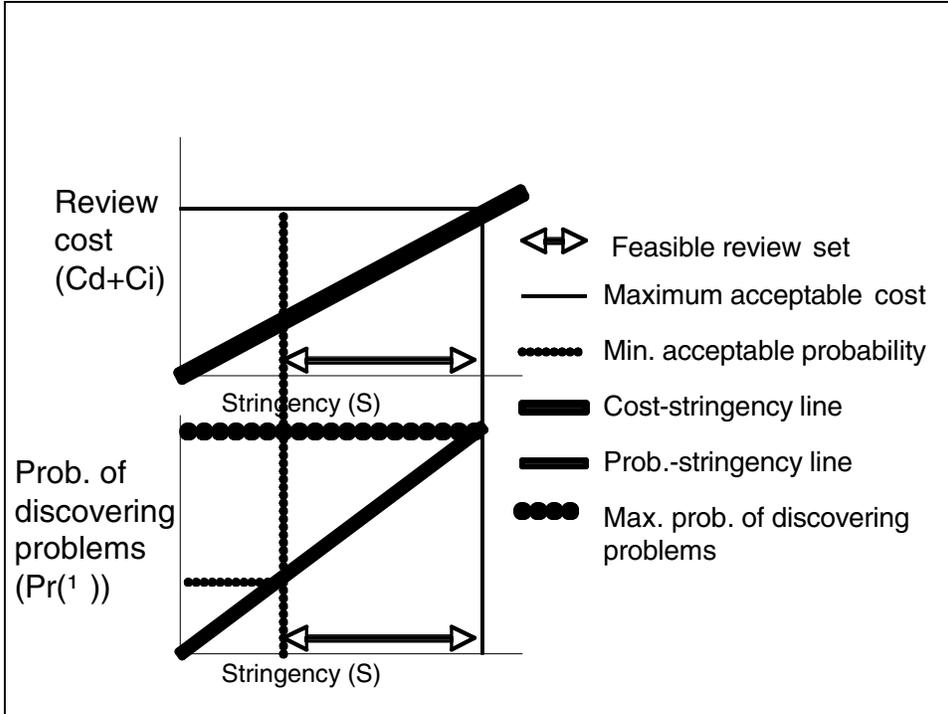


Figure 3: Why Judicial Review Fails

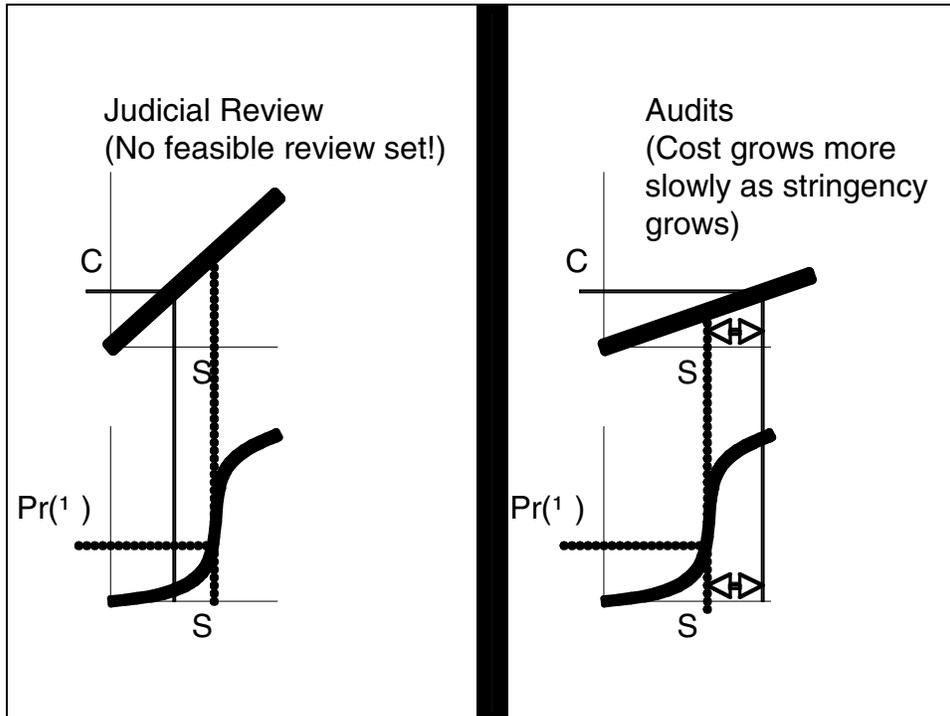


Figure 4: Methodology Used in a Stratified Random Sample of GAO and IG Office Reports, January 2002 – January 2005

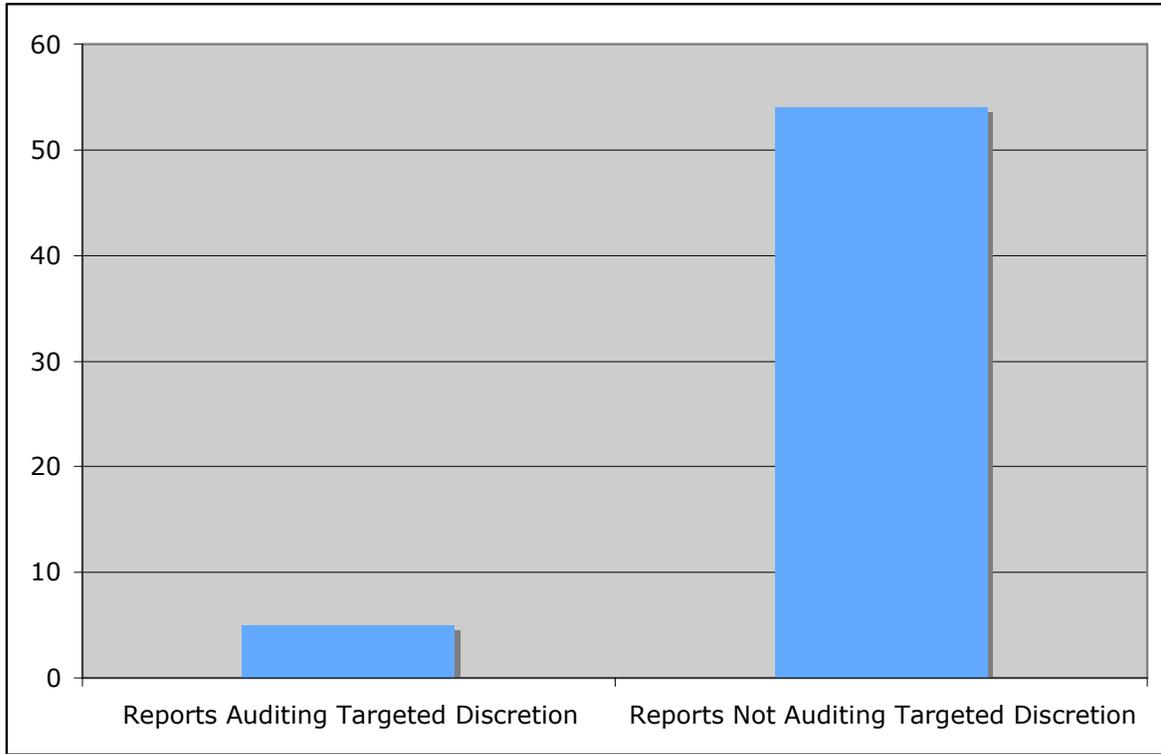


Figure 5: New York Times Reports on Audit Bureaucracies, January 2002-January 2005

