

[v 3.99 draft; Please don't quote without the permission of the author, his former roommate's two cousins, and at least one dog]

AUDITING EXECUTIVE DISCRETION

Mariano-Florentino Cuéllar*

Executive branch officials routinely make thousands of decisions affecting public security and welfare. While it is rare that such discretionary decisions are entirely immune from some kind of judicial review, courts' role is often so circumscribed or deferential that in some domains the probability of uncovering problems through such review almost certainly falls close to zero. The resulting amount of executive discretion carries considerable risks along with rewards. Some discretionary decisions undoubtedly benefit from the speed and flexibility that results from limiting judicial review. Yet judicial review's evisceration as a tool to restrain certain forms of discretion also makes it easier for 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. Reliance on judicial review to generate information about executive discretion makes it difficult 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 potential case in a particular class, making it difficult to increase the stringency of review in some policy domains without making the costs allegedly prohibitive. When deciding how stringently to review a discretionary decision – whether a prosecutorial charging decision, an administrative compliance order, or an enemy combatant designation – judges almost invariably mull the potential consequences of their choice on all future executive decisions of that kind. As a conceptual alternative, this article develops a framework akin to that employed by courts engaged in sample adjudication for class action and government fraud cases. It relies on the possibility of systematically auditing samples of discretionary decisions and making those results public. Although the efficacy of such a system depends on the political context and details of its institutional design, audits have the potential to sever the connection between the perceived costs of encroaching on discretion and the stringency of review. They also avoid the potentially distorted picture of bureaucratic activity created by a litigation-driven process. Despite their potential value, such audits are nonetheless almost never undertaken by existing federal audit bureaucracies (the Government Accountability Office and the department-specific Inspector General Offices), nor does the legislature seem to conduct them in connection with oversight hearings. I conclude by discussing the political and bureaucratic dynamics working against these audits and suggesting how they may be weakened.

INTRODUCTION	2
I. THE LOGIC OF AUDITING EXECUTIVE DISCRETION	7
a. <i>How Executive Discretion is Pervasive</i>	9
i. <i>Overview and Definitions</i>	9
ii. <i>Examples</i>	11
1. <i>Asset Freezing</i>	11
2. <i>No-Fly Lists</i>	13
3. <i>Environmental and Occupational Safety</i> <i>Administrative Compliance Orders</i>	15
b. <i>How Discretion is Routinely Managed Through Variations in the</i> <i>Stringency of Court Review</i>	16

* 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, Daniel Sutherland, Mark Kelman, Pam Karlan, Barbara Fried, Deborah Hensler, Rachel Barkow, Mark Lemley, Jayashri Srikantiah, Mitch Polinsky, Lawrence Friedman, Michael Barr, Sallyanne Payton, Sally Katzen, and Jeff Strnad, Chris Elmendorf and feedback from participants at a University of Michigan Legal Theory Workshop and a Stanford Law and Economics “Free Lunch.” I am especially indebted to Richard Craswell, Peter Strauss and Alan Morrison. If an audit reveals any errors or omissions they are my responsibility.

c.	<i>How Audits Can Substitute For, or Supplement, Judicial Review</i>	18
i.	<i>Envisioning the Institutional Design of Audits</i>	19
ii.	<i>Trading Off Breadth Against Depth Assuming Marginal Cost Analysis</i>	23
iii.	<i>Alternative Assumptions: Separate Informational and Cost Aspirations</i>	24
d.	<i>Why Traditional Judicial Review Suffers From Limitations</i>	28
i.	<i>Reviewing Rationales for Departing From a Limited-Discretion Baseline</i>	28
ii.	<i>Why Bureaucracies Granted Discretion Fail (to Learn)</i>	30
e.	<i>How Audits Can Ameliorate the Limitations of Traditional Judicial Review</i>	33
II.	WHY THE INSTITUTIONAL DESIGN PROBLEMS INHERENT IN AUDITING EXECUTIVE DISCRETION ARE MANAGEABLE	37
a.	<i>Audits Can Be Adapted to Address Multiple Forms of Discretion</i>	37
b.	<i>Most of the Audit -Related Institutional Design Problems Have Plausible Solutions</i>	39
III.	WHY REFORMS WOULD CONTRAST SHARPLY WITH THE STATUS QUO	44
a.	<i>Existing Audit Bureaucracies Do Not Ordinarily Audit Executive Discretion</i>	44
b.	<i>Neither Do Legislators</i>	47
IV.	WHY AUDITS FACE POLITICAL AND ORGANIZATIONAL CONSTRAINTS	
a.	<i>Political Actors Have Polarizing Incentives</i>	49
b.	<i>Institutional Inertia Locks In Conceptions of Adjudication</i>	51
c.	<i>Narrow Conceptions of Auditor Mission Persist</i>	52
CONCLUSION	55

INTRODUCTION

The problem is a familiar one. No legal system can ever wholly vanquish discretion – and even if it were possible, it would be madness to wring all human judgment from the application of legal rules and standards. But too much discretion breeds its own kind of madness. To manage this predicament, our government subjects many decisions to elaborate procedural constraints, but in other domains it ostensibly seeks to harness the value of human judgment by leaving public officials with considerable discretion. On a typical day, Labor Department officials decide what plants to inspect for occupational safety violations with little or no external review. Prosecutors decide whom to indict. Treasury officials decide whether to freeze the assets of a charity because of alleged links to terrorism. Homeland security inspectors decide whether a Namibian woman will be turned away at a port of entry without being allowed to plead her case for asylum, and whose name is placed on a government “no fly” list. In short, despite an often-mentioned social commitment to judicial restraints on public power,¹ laws routinely create and protect discretion instead of restraining it.²

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

This absence of restraints, in turn, raises the pressing question of how public organizations with such considerable discretion will (be forced to) learn from their mistakes, and how they will be policed against abusing their legal powers. While judicial review of government action is often considered a central tool in preventing mistakes or abuses,³ the benefits of maintaining discretion have led to a vigorous doctrinal and policy debate about the proper stringency of such review.⁴

This article challenges the terms of that debate. The argument demonstrates how the paradigm of judicial review, despite its enduring value, sometimes ill-serves the goals of helping bureaucratic organizations learn from their failures and avoid political

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).

² See *infra* notes ____.

³ 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."); Jerry L. Mashaw, *Reinventing Government and Regulatory Reform: Studies in the Neglect and Abuse of Administrative Law*, 57 U. PITT. L. REV. 405, 414 (1996) ("The long-term commitment of American administrative law has been to assure that administrative discretion is structured, checked, and balanced. Administrative efficacy must be weighed against demands for liberty and legality, as well as political accountability"); 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). Understandably, judicial review is considered important both because of its role in generating information about, and promoting the accountability of, executive authorities, and because of its direct remedial role. Regarding "accountability" see, e.g., Jeffrey J. Rachlinski and Cynthia R. Farina, *Cognitive Psychology and Optimal Government Design*, 87 CORNELL L. REV. 549, 587-89 (2002); Mark Seidenfeld, *Cognitive Loafing, Social Conformity, and Judicial Review of Agency Rulemaking*, 87 CORNELL L. REV. 486, 509 (2002); Michael Asimow, *The Scope of Judicial Review of Decisions of California Administrative Agencies*, 42 UCLA L. REV. 1157, 1239 (1995); William Christian, Note, *Normalization as a Goal: The Americans With Disabilities Act and Individuals With Mental Retardation*, 73 TEX. L. REV. 409, 417 (1994). The importance of providing a hearing and a remedy for government actions affecting a protected interest, meanwhile, is firmly lodged at the core of modern procedural due process doctrine. See *Mathews v. Eldridge*, 424 U.S. 319 (1974).

⁴ Voices on one side of the debate emphatically insist on greater opportunities for highly-stringent judicial review of executive branch actions. See, e.g., 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; Nicole Nice-Petersen, Note, *Justice for the "Designated": The Process That is Due to Alleged U.S. Financiers of Terrorism*, 93 GEO. L.J. 1387, 1419 (2005) ("The courts should not hesitate to act solely because those stripped of their rights are accused terrorists"). Similarly emphatic voices take the position in equipoise. See, e.g., Eric A. Posner and Adrian Vermeule, *Accommodating Emergencies*, 56 STAN. L. REV. 605, 644 (2003) ("Judicial scrutiny can only interfere with forceful executive action."); 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 Part III.

pressures endangering their missions.⁵ Because complex public bureaucracies are increasingly (and perhaps inevitably) the custodians of discretionary legal authority that can be abused,⁶ the problems arise both in national security and domestic regulatory contexts, domains that have been traditionally treated separately but increasingly blur.⁷ The article then shows how a government agency can perform quasi-judicial *audits* of discretionary decisions, akin to the sample audits occasionally employed by courts in class actions and government fraud cases.⁸ If an auditing agency overcomes the relevant political barriers and conceptual challenges,⁹ it can fill crucial gaps left by existing mechanisms to generate information about executive branch performance. Even if those barriers prove too onerous to surmount, the conceptual work associated with designing an audit system can serve as a thought experiment, clarifying the murkiness of conventional defenses of broad executive discretion, and shedding light on why such claims nonetheless resonate in the legal and political arena.

The argument begins by assuming, for expositional purposes, that judicial review is the only institutional mechanism generating information about discretionary executive decisions directly affecting individuals or groups, such as whether to decide that an individual lacking documents at the border has shown enough “credible fear” to be allowed to apply for asylum. The picture of executive discretion that emerges in such a world will inevitably depend on the structural features of judicial review. That picture will reflect, for instance, the courts’ tendency to balance the potential benefits and costs of discretion by routinely applying differing degrees of stringency when reviewing executive decisions. Suppose, for instance, that the issue is the fate of individuals that executive authorities designate as enemy combatants. Whether on their own or in accord with legislative commands, 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.¹⁰ Greater stringency of review presumably reduces the

⁵ Although this article does not directly address judicial review’s role as provider of individual remedies, the argument developed here is nonetheless relevant to the provision of remedies by either courts or political actors. *See supra* notes ___ (discussing the relevance of the argument to court review of individual cases); ___ (discussing the implications of the argument for how political actors deliver remedies to aggrieved individuals or groups).

⁶ The focus here is primarily on the type of discretion, such as that vested in a prosecutorial authority, to impose costs on discrete individuals or groups with minimal judicial intervention. *Cf. Heckler v. Chaney*, 470 U.S. 821 (1985)(in the absence of specific statutory requirement to the contrary, regulatory agency’s decision not to exercise authority in a particular context where such authority could be exercised is committed to agency discretion) *See also infra* note___ and accompanying text.

⁷ *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.”).

⁸ *See infra* note ___.

⁹ For a discussion of the proverbial “guarding the guardians” problem, see *infra* notes 190-191 and accompanying text.

¹⁰ *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 do not mean to

probability that someone would be improperly labeled an enemy combatant. At the same time, greater stringency allegedly increases the resources that society must expend on the review process and that the executive branch must expend defending its decision.¹¹ 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. In response, courts tend to vary the stringency of review governing a given pool of potential cases.

But this strategy of applying a given degree of review stringency across the board entails its own costs. By forging rules applying to every case in a particular class, courts and legislators can impose dramatic limits on society's ability to learn how executive discretion is used.¹² Increasing the stringency of review for a single decision appears difficult, if not impossible, without sharply (and, it is sometimes asserted, dangerously) increasing costly burdens on courts and the government. Even if one assumes (implausibly) that existing rules governing stringency of review reflects a careful analysis of marginal costs and benefits, existing limits on review stringency almost certainly augur problems for society's ability to learn how discretion is used.

Government regulators and private employers, in contrast, routinely manage to avoid such traps. Instead of reviewing an entire population's behavior, they obtain samples of it. Insurance companies examine a subset of closed files to assess the quality of payout determinations.¹³ Government agents select a subset of plants to inspect or accounting records to scrutinize.¹⁴ The tactic can be easily adapted by a court-supervised or independent authority to generate information about public bureaucracies. The defining feature of this technique is its rejection of an implicit assumption that a given degree of review stringency should be applied to all cases in a class. Despite their

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 the very 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.

¹¹ 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...”).

¹² Moralistic intuitions about the importance horizontal equity combine with the content of legal doctrines such as *stare decisis* to complicate the possibility of using differing degrees of stringency to review cases in the same class. See *infra* Part I.b.

¹³ See *infra* Part I.c.

¹⁴ See *id.*

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

Nor would audits merely duplicate, in a different form, what judicial review could already achieve. Under quite reasonable assumptions, deferential judicial review may be worse from a prescriptive standpoint 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.¹⁵ 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. The information on bureaucratic performance generated by audits could even dynamically influence courts deciding how stringently to review government action, or whether such action comports with procedural due process norms.¹⁶ Ultimately, audits may also exert an impact on the political context shaping the allocation of power to government. In principle, that context should reflect an accountability-power trade-off, where political audiences may prove 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.

Part III then relaxes the assumption that judicial review is the only institutional mechanism generating information about executive discretion. It examines the behavior of two sets of actors capable of producing the sort of accountability-enhancing information that intrusive judicial review can. The Government Accountability Office (GAO) and the Inspectors General (IG) offices, which might be termed "audit bureaucracies," possess a broad mandate to audit federal government activities and produce information about how laws are implemented. Although they operate in a complex political environment shaped by the legislature and the executive branch, they are not subject to the constraints that ostensibly lead courts to fashion stringency standards applying across the board to all potential cases in a particular class. Similarly,

¹⁵ Of course, the mere creation of some auditing system does not automatically solve organizational learning and accountability problems involving the law. As noted 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 the role of audits and the pitfalls in designing them, see MICHAEL POWER, *THE AUDIT SOCIETY: RITUALS OF VERIFICATION* (1999).

¹⁶ See *infra* Part II.b (discussing the potential for dynamic interaction between audits and judicial review).

the legislature itself can use its investigative powers directly to generate information about how executive discretion is exercised.

Yet audits of executive discretion generally are not a feature of modern governance. Amidst the swirl of budget votes in Congress, committee hearings, GAO investigations of FBI computers, and IG reports on immigration policy, neither legislatures nor the audit bureaucracies focus on systematically auditing executive discretion.¹⁷ As Part IV indicates, the relative absence of sampling techniques may reflect conceptual blurring of the direct remedy and information-producing remedies that bedevil courts, and probably lead courts to under-use the sampling methodologies that some judges have cautiously deployed in class actions and government fraud cases. In addition, legislators and organized interests, like the executive branch itself, may lack incentives to deploy audits or analogous sampling methodologies. This pattern of neglect predictably affects how legislatures bargain over executive power and review the consequences of those bargains. It also affects how legislative goals percolate through the audit bureaucracies, drawing their attention to procurement fraud and similarly tangible examples of waste. Second, to the extent that the audit bureaucracies retain some autonomy to allocate their resources, they appear to remain in the thrall of their initial role as financial auditors and the more-recently acquired role of auditor of government performance that nonetheless generally fails to encompass sampling. While this picture is not immediately encouraging, it does suggest that as audit bureaucracies expand their autonomy, they may be able to entice a constituency to value audits of executive discretion if agency leaders choose to pursue such a goal.

Even in the face of such efforts to enhance the organizational autonomy of the audit bureaucracies, the preceding factors may nonetheless continue locking in suboptimal institutional responses to serious legal problems. This case study and thought experiment therefore aims to highlight three crucial challenges that follow from that sort of lock-in: (1) the importance of recognizing the inherent limitations of traditional 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. In response to those challenges, this article seeks to broaden the scope of potential solutions and shed light on the forces that shape public perceptions of whether those solutions even exist.

I.

THE LOGIC OF AUDITING EXECUTIVE DISCRETION

It is commonplace for scholars and policymakers to extol the role of judicial institutions in ensuring that rights are respected, that agencies do not exceed their legal powers, and that the law is correctly applied.¹⁸ Legislators have bolstered the judicial

¹⁷ As Part I.c explains, this means that the audit bureaucracies generally appear not to: (1) take random samples of decisions (2) in an important legal domain where decisions directly affect individuals or groups, (3) assess those decisions in accordance with a defensible standard and announce the results to the public.

¹⁸ See, e.g., Peter M. Shane, *Interbranch Accountability in State Government and the Constitutional Requirement of Judicial Independence*, 61 LAW & CONTEMP. PROBS. 21 (1998), and sources cited *supra* at note 1.

role over the years by enacting procedural controls, such as those contained in the Administrative Procedure Act, enhancing their ability to assert control over the bureaucracy and limiting the scope of executive discretion.¹⁹

Despite these developments, court review of executive decisions is sometimes exceedingly modest, or entirely missing.²⁰ The pattern recurs in domains involving both national security and more traditional domestic administrative action. Sometimes legislators explicitly confer enormous discretion, 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).²¹ In other cases, executive agencies have discretionary powers because prevailing statutory interpretations and constitutional provisions imply the existence of such power. Courts have recursively found that prosecutors harbor an inherent power to choose whom to charge with few (if any) judicially-imposed constraints.²² In still other cases agencies have discretionary power because courts and other external observers 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.²³ The resulting degree of executive discretion lets agencies act swiftly in domains where they hold responsibility, learn to deal with unfamiliar situations, and leverage their expertise.²⁴

Yet discretion's two-edged nature constantly reiterates the prescriptive question of how best to render accountable the executive bureaucracies wielding such power. This Part explains why audits of discretionary decisions should play a prominent role in answering that question.

A. *How Executive Discretion is Pervasive*

i. *Overview and Definitions*

¹⁹ See generally, McNollgast, *The Political Origins of the Administrative Procedure Act*, 15 J.L.ECON. & ORG. 180 (1999); Peter L. Strauss, *Changing Times: The APA at Fifty*, 63 U. CHI. L. REV. 1389 (1996).

²⁰ Even when all or most alternative review is precluded, 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 interfere with some of the more valuable properties associated with executive discretion. See generally PETER SCHUCK, *SUING GOVERNMENT: CITIZEN REMEDIES FOR OFFICIAL WRONGS* (1983).

²¹ *Webster*, 486 U.S. at 602-04.

²² *United States v. Armstrong*, 517 U.S. 456 (1996). As 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." Davis, *supra* note 1, at 209 n.21.

²³ See *Hamdi v. Rumsfeld*, 542 U.S. 507 (2004).

²⁴ 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).

For present purposes, discretion can be defined as the extent of legal flexibility to use government power vested in executive branch officials – including, but not limited to, personnel, budgets, information, and legally-sanctioned coercive authority to affect the world.²⁵ Distinctions in the amount of executive discretion are relative, not absolute. Government officials exercise a certain measure of discretion virtually every time they do something. Though government actions are rarely purely discretionary, neither is discretion ever entirely absent. 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.

Courts and commentators have long acknowledged the importance of some deviations from that baseline. Thus, while the classic case of *Citizens to Preserve Overton Park v. Volpe* firmly establishes close judicial scrutiny of a discretionary decision as a presumptive means of policing executive decisions,²⁶ other cases among the administrative law canon make an equally compelling case for a greater measure of discretion. In some cases the Court famously recognized the existence of domains of government authority best left, for structural reasons, entirely to the political branches.²⁷ A somewhat different form of discretion is the subject of *Heckler v. Chaney*,²⁸ where the Court considered the appropriate degree of judicial scrutiny of a regulatory agency’s discretionary enforcement decisions, and emphasized the value of limiting judicial intervention in such matters by analogizing regulatory enforcement discretion to prosecutorial discretion in the criminal context.²⁹ It is this form of targeted discretion to directly affect individuals and groups that most resembles the “executive discretion” with which this article is concerned. As many commentators have acknowledged, preserving the type of prosecutorial discretion at issue in *Heckler* from judicial intervention raises the risk that executive authorities will behave in an arbitrary fashion.³⁰

When lawyers advocating on behalf of executive power extol the value of such *Heckler*-type executive discretion, they nonetheless tend to implicitly accept a baseline

²⁵ One could imagine, in contrast, a different type of discretion, understood to require by definition – perhaps for political reasons – the absence of review. Efforts to subject this sort of discretion, associated (for instance) with the political question doctrine, to review would be (by definition) ill-advised. Nonetheless, one of the goals of the discussion that follows is to highlight the value of being more explicit about *why* certain types of discretionary decisions should be entirely insulated from review.

²⁶ 401 U.S. 402 (1971).

²⁷ See, e.g., *Marbury v. Madison*, 5 U.S. (1 Cranch) 137, 170, 2 L.Ed. 60 (1803). As I acknowledge below, perhaps there are indeed some contexts where discretion does (and ought to) mean no review -- from courts, auditors, or anyone else. My suggestion is simply that we ought to proceed with great caution in considering the rationales for that kind of discretion. In particular, we ought to recognize that the rationales for limiting or even barring judicial review in some contexts do not necessarily speak to the question of whether we should also bar an alternative form of review. Put differently, perhaps one might envision an array of review mechanisms to address the fact that the two kinds of discretion are on a continuum, and at intermediate places on the continuum we might do well to use audits (or some other system of review) instead of leaving the executive unrestrained.

²⁸ 470 U.S. at 821.

²⁹ The analogy bolstered the Court’s case for limiting judicial interference in the *Heckler* context (involving the Food and Drug Administration) because the perception was already so deeply-rooted among courts that judicial regulation of prosecutorial discretion would unduly burden the administration of justice. Cf. *Bordenkircher v. Hayes*, 434 U.S. 357, 664 (1978).

³⁰ See *infra* note ____.

state of the world where courts play a significant role in reviewing government action.³¹ Such recognition of the value of judicial supervision is a familiar one in the United States and in most other developed nations (and many developing ones).³² In criminal prosecutions, voting rights cases, and labor law injunctions, for example, the completion of some action of the executive branch (such as subjecting someone to the detriments associated with being convicted of a crime) is conditioned on judicial approval. 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 fail in persuading the legislature to water down the substantive standard that applies, advocates of discretion are left to mount a vigorous case before a court that is quite persistently unwilling to simply defer to executive discretion. Even when such review does not occur in advance of government action, 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 subsequently invalidated.³³ This implies that we can measure the *benefits* of executive discretion against a baseline of relatively intrusive judicial review.

ii. Examples

³¹ 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).

³² 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).

³³ 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%).

To better assess the potential benefits and costs of departing from a baseline of intrusive judicial review, we should first analyze how executive authorities may use their discretion in a few specific contexts arising in both national security and domestic regulatory enforcement. Consider, for instance, the following examples.

1. Asset Freezing – Under current law, government officials have powerful tools to regulate the economic activities of foreign persons. One such tool, involving designations under the International Emergency Economic Powers Act (or “IEEPA”),³⁴ 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 something in which a foreign country or national has an interest.³⁵ This constraint turns out not to be 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 American organization has foreign beneficiaries may be enough, in fact, for a court to say that it has a “foreign interest.”³⁶ Second, the president must use this power only during an emergency.³⁷ This is not much of a limitation, either. The “unusual and extraordinary threat” giving rise to the emergency must have its source partly outside the United States.³⁸ It must pose a threat to the “national security, foreign policy, or economy of the United States.”³⁹ Given the combined effect of this expansive language and traditional judicial deference on matters of national security and foreign affairs, presidents have found it relatively easy to declare emergencies under the law (about ten of which are currently in effect).⁴⁰ Courts have yet to find, under the terms of IEEPA, that a supposed emergency does not exist.

In a series of executive orders, the President has delegated much of his 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.”⁴¹ Once a group becomes a specially designated terrorist organization, it loses control over its fate. The impact of OFAC’s orders is to block the organization’s funds, regardless of where in the financial system they happen to be. Each violation of the blocking orders can trigger a separate civil fine of up to \$10,000, and willful violators are subject to criminal penalties, including up to ten years’ imprisonment and fines of up to \$50,000 per violation.⁴² A similar designation also triggers severe criminal penalties punishing individuals for providing “material

³⁴ 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).

³⁵ *Id.*

³⁶ *See* Global Relief Fdn. V. O’Neill, 315 F.3d 748 (7th Cir. 2002).

³⁷ *See* IEEPA, *supra* note 70.

³⁸ *See* Holy Land Fdn. v. Ashcroft, 333 F.3d 156, 159 (D.C. Cir. 2003).

³⁹ *Id.*

⁴⁰ *See generally* Jason Luong, Note, *Forcing Constraint: The Case for Amending the International Emergency Economic Powers Act*, 78 TEX. L. REV. 1181 (2000).

⁴¹ 333 F.3d at 160.

⁴² 50 U.S.C. § 1705 (2004).

support” (including funds and in-kind economic contributions such as lodging) to designated terrorist organizations.⁴³

Take a closer look at how a court reviews the government’s designations. In the recent *Holy Land Foundation* case, the State and Treasury departments used their delegated presidential IEEPA powers to freeze the assets of the Holy Land Foundation.⁴⁴ Court review of the blocking order considered whether it was “arbitrary and capricious” under the terms of the Administrative Procedure Act. But given the national security context of the decision, the reviewing district and appellate courts also interpreted the relevant law to require a highly deferential form of review. The District Court, for example, emphasized the limited scope of its role. 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.⁴⁵

The D.C. Circuit upheld this determination on appeal.⁴⁶ In doing so, the court legitimized a review process that arguably renders administrable the federal government’s web of emergency economic regulatory powers. That process might also strike some observers as particularly thorough. But whatever one’s views about the former issue, the latter 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 [instead only allows the court] to review the agency’s record to determine whether the agency’s decision was supported by a rational basis.”⁴⁷ That record may be a tremendously accurate complication of the government’s evidence. Or it may be patently misleading. Nothing requires the government to report evidence tending to cast doubt on

⁴³ See *Humanitarian Law Project v. Reno*, 205 F.3d 1130 (9th Cir. 2000).

⁴⁴ 333 F.3d at 159-60.

⁴⁵ 333 F.3d at 161.

⁴⁶ 333 F.3d at 161-63.

⁴⁷ See *Holy Land*, 219 F.Supp.2d at 67.

its contentions.⁴⁸ 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. A court will vacate the designation if the record in question turns out to be an empty folder. 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 designation look totally arbitrary. Yet the record itself is based on decisions that are essentially immune from review.⁴⁹

2. *No-Fly Lists* –When passengers check in at American airports, their names generally are compared to those on a list provided to airlines by the Transportation Security Administration (TSA).⁵⁰ Predictably enough, the point of the list is to thwart the (travel) plans of those who might prove dangerous on board a flight.⁵¹ Although TSA initially denied the existence of this list (technically an element of a system known as CAPPS I), subsequent disclosures establish that TSA considers an important component of aviation security even in light of recent efforts to supplement it with a more elaborate system.⁵²

Yet TSA and the airlines have run into a spate of difficulties using the list. Sometimes people whose names are merely similar to those on the list are detained. Some individuals 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

⁴⁸ 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.

⁴⁹ Manipulation of the record, moreover, need not 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)].

⁵⁰ See Government Accountability Office, *Aviation Security: Computer-Assisted Passenger Prescreening System Faces Significant Implementation Challenges*, GAO 04-385 (Feb. 2004) (“GAO Report”).

⁵¹ A subsidiary purpose may also be to promote the questioning and apprehension of suspicious individuals who may be sought by law enforcement authorities.

⁵² See generally *Green v. TSA*, 351 F.Supp.2d 1119 (W.D. Wash. 2005). See also GAO Report, *supra* note 86, 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). 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. See *EPIC v. Dep’t of Homeland Security*, ___ F.Supp.2d ___, 2005 WL 1745303 (July 25, 2005). 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). 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. 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. See also GAO Report, *supra* note ___, at 9.

to extensive questioning.⁵³ Even 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 appears to exist.⁵⁴ Instead, people who plainly should not be on the list are bewildered to learn that they cannot even find out *how* their name appeared on the list, let alone what must be done to remove it.⁵⁵ A member of the Air Force reporting for duty found his name on the list. A mother checking in at Dulles Airport for a flight to Italy discovered her gurgling nine-month old son's name was on the list, which prevented the desk agent from printing out a boarding pass.⁵⁶ 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, particularly given the apparent absence of any reliable, consistent review mechanism to ensure the names on the list belong there.⁵⁷

Reasonable people can differ with respect to how much discretion TSA should have in deciding what information to consider in administering the list. What is harder to deny is the absence of any specific statutory provision to resolve disputes about the propriety of placing a person's name on the list.⁵⁸ Efforts to persuade a court that that the TSA's decision to place a name on the list was "arbitrary and capricious," moreover, would encounter the same problems associated with the blocking of assets, where courts have apply an exceedingly deferential standard that essentially fails to encompass review of how the agency obtained the information that allegedly supports its determination.⁵⁹ 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.⁶⁰ Regardless of the outcome, the no-fly list 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

⁵³ See GAO Report, *supra* note __, at 13.

⁵⁴ See *id.* (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).

⁵⁵ 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. Individuals who write TSA may also obtain a letter that is supposed to mitigate the impact of being found on the list. But presentation of the letter to airline or TSA authorities at the airport is no guarantee that a person whose name appears on the list will be allowed to travel. See Drees, *supra* note __.

⁵⁶ Caroline Drees, *U.S. No-Fly List Vexes Travelers From Babies On Up*, REUTERS ONLINE, (Dec. 19, 2005), avail. at http://today.reuters.com/news/newsarticle.aspx?type=domesticNews&storyid=2005-12-15T151246Z_01_SIB554549_RTRUKOC_0_US-SECURITY-NOFLY.xml&rpc=22 (last accessed Dec. 19, 2005)(observing how the suspected nine-month old terrorist "sat there and gurgled," and noting that the list apparently includes Senator Edward Kennedy and Representative Don Young among others).

⁵⁷ See GAO Report, *supra* note 86, at 13.

⁵⁸ Persons may *complain* to an ombudsperson or directly to the agency, but there is no administrative mechanism to resolve such complaints.

⁵⁹ An arbitrary and capricious claim would also confront an additional hurdle in this context,

⁶⁰ Such a lawsuit is currently pending.

violate the interest-balancing test rooted in *Mathews v. Eldridge*.⁶¹ 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.

3. *Environmental and Occupational Safety Administrative Compliance Orders* – By subjecting individuals to severe practical and reputational consequences, harsh criminal indictments may operate as discretionary sanctions. But individuals and organizations tend to face formal punishments in the criminal justice system only after they are convicted or admit their guilt. Statutes creating major regulatory programs reflect a different premise. Many such laws allow regulators to levy fines or issue orders restricting certain activities with more limited court intervention. Although they vary in the relevant legal standard or the size of the maximum fine, those orders can have an effect before judicial intervention. Even after that intervention, it is not 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.

For instance, when the Occupational Safety and Health Administration (OSHA) or its contractors has some reason to think businesses are violating their general duty to provide a safe working environment, the agency can issue abatement orders and citations.⁶² 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.⁶³ 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 does not allow EPA to impose fines or other penalties directly, the order triggers provisions imposing civil or criminal penalties for violation of the order.⁶⁴ 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.⁶⁵ 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 has found ACOs not to be final agency actions, thereby rendering them unreviewable and raising the due process problem.⁶⁶ The Supreme Court declined to review this case, and instead – in a separate case – upheld a Ninth Circuit

⁶¹ 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).

⁶² See 29 U.S.C. §§664(a), 666(a), 666(b), and 666(c).

⁶³ Clean Air Act (“the Act”), 42 U.S.C. §§ 7401-7671q.

⁶⁴ 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).

⁶⁵ See 42 U.S.C. § 7413(b)(2).

⁶⁶ See *TVA v. Whitman*, 226 F.3d 1236 (11th Cir. 2003).

opinion holding ACOs to be final agency actions and reviewing them under the “arbitrary and capricious” standard.⁶⁷

Both contexts leave the agencies considerable discretion to impose compliance orders. OSHA obviously has it when it issues citations and abatement orders, some of which are not challenged subsequently. Even if arbitrary and capricious review is not as deferential in this context as it is with asset freezes, it still leaves the court applying a fairly deferential standard of review to a decision that can be based on “any available information.” It is 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 is 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, more exhaustive 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.

*

*

*

The preceding examples demonstrate the limits of judicial scrutiny for certain vital government functions. Indeed, even the availability of substantive review can conceal vast reservoirs of bureaucratic flexibility. A host of other domains – varying in the availability of formal review but not in the fact that they leave authorities with discretion – pose similar problems. Such domains involve, among others, the impact of prosecutors’ charging decisions on suspects, the meager judicial scrutiny of discretion vested government contractors engaged in quasi-official functions,⁶⁸ the enormous power federal officials wield (even in the wake of recent Supreme Court decisions on the subject) in designating enemy combatants,⁶⁹ and even in the myriad decisions governing federal procurement predicated on the exercise of government officials’ legal discretion.⁷⁰ How, then, should courts and legislatures react if they consider the existing degree of external scrutiny to be inadequate?

B. How Discretion Is Routinely Managed Through Variations in the Stringency of Court Review

The answer is not, obviously, to maximally limit discretion, something courts and legislatures are understandably loath to do. Instead, judges and lawmakers tend to manage the costs and benefits of discretion by varying the stringency of review that is supposed to apply to the actions of the executive or her agent. Whether courts are driven to do this by anodyne judicial prudence or rigid legislative mandates, they review some

⁶⁷ 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).

⁶⁸ See generally Gillian Metzger, *Privatization As Delegation*, 103 COLUM. L. REV. 1367 (2003). See also Steven L. Schooner, *Contractor Atrocities at Abu Ghraib: Compromised Accountability in a Streamlined, Outsourced Government*, 16 STAN. L. & POL’Y REV. 549 (2005).

⁶⁹ See *Hamdi*, 542 U.S. at 518-19.

⁷⁰ See Jody Freeman, *The Contracting State*, 28 FLA. ST. U. L. REV. 155, 165 (2000) (noting that, despite the “highly technocratic approach to contract design” prevalent in federal procurement law, the existing framework is “too limited to address the much more substantial issues that arise” in some contracts).

decisions more stringently, and others less so.⁷¹ Even decisions putatively subject to the same standard of review, such as the familiar arbitrary and capricious touchstone enshrined in the Administrative Procedure Act, may end up being reviewed with different degrees of stringency. The distinctions presumably reflect courts' judgments about when the costs of added scrutiny are justified.⁷² Thus, judges and scholars generally take arbitrary and capricious review to mean one thing (milder review) for a typical informal adjudication, such as deciding whether a vehicle fits standards permitting to enter a national forest, and another (more stringent review) when courts are reviewing an intricate regulatory rule governing the licensing of nuclear reactors.⁷³

Stringency of review is what distinguishes these two different versions of the arbitrary and capricious standard, and more generally, what differentiates government decisions that receive greater judicial scrutiny from those that get less. The term is meant to serve as an abbreviated reference to the mixture of doctrines governing such distinctions in the strictness of review applied to an agency's factual or prescriptive conclusions in a given decision. Stringency includes, 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 can imply 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.

Thus, when a court determines that a six-page declaration from an official ensconced beneath layers of the Defense Department bureaucracy is enough reason to detain someone for an indefinite period of time, it is being more deferential.⁷⁴ When a court decides such justification is insufficient, 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.⁷⁵ 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 auguring to provide review that is even more

⁷¹ For cases, *see supra* note 6.

⁷² Indeed, casual observers may be forgiven for assuming (heroically) that courts (or legislatures, when they directly impose limits on review) are balancing the marginal costs and benefits of greater stringency of review. *See infra* Part I.c.ii for a discussion.

⁷³ *See, e.g.,* AMERICAN BAR ASSOCIATION - SECTION OF ADMINISTRATIVE LAW AND REGULATORY PRACTICE, A BLACKLETTER STATEMENT OF FEDERAL ADMINISTRATIVE LAW 34 (2004). The Blackletter Statement provides a cogent and revealing synthesis:

The court may set aside an agency action as an abuse of discretion... on any of several grounds. In practice, application of these grounds varies according to the nature and magnitude of the agency action. *Thus, a court will typically apply the criteria rigorously during judicial review of high-stakes rulemaking proceedings (a practice commonly termed "hard look" review), but much more leniently when reviewing a routine, uncomplicated action.* [Emphasis added].

⁷⁴ *See Hamdi*, 316 F.3d 450, 473 (4th Cir. 2003).

⁷⁵ *See Hamdi*, 542 U.S. at 519.

stringent.⁷⁶ Though specific cases may involve different doctrinal bases and legal subtleties, the important distinction is not in the specific doctrines or procedures involved. It is in the fact that under a standard of greater deference, whether caused by legislative enactments or subsequent judicial interpretation, some executive bureaucracy retains greater power to decide how to use its discretion.

Decisions about review stringency, moreover, follow a certain convention. When deciding how much of that power to let executive authorities keep, courts and legislators tend to implicitly assume 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.⁷⁷ In fact courts treat horizontal equity as an important value, where deviations must be defended.⁷⁸ The same goes for virtually all the legislative mandates that courts implement. The move to privilege horizontal equity in judicial review is the essence of traditional judicial review.⁷⁹ Its rationale may be grounded in an appreciation for *stare decisis*, or perhaps resides in an inflated conception of judicial power to ensure that like cases are treated similarly. Predictably, the desire for horizontal equity in standards governing the stringency of review renders troubling (at least in the eyes of many principled observers) the prospect of increasing the stringency of review in a particular case. A decision to increase review stringency in a single case is taken to cast a long shadow on all similar decisions in the relevant pool of cases, and prohibitively increasing the associated costs in terms of the direct burdens of review and the forgone benefits of discretion.

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

Suppose now that those allegedly prohibitive costs, or their related administrative complications, preclude extensive supervision of executive discretion through judicial review. To fill the gap, some observers may rely on heroic conceptions of the political process, or the media. But each of these mechanisms depends in large measure on information about the nature and quality of executive decisionmaking. Politicians and

⁷⁶ No doubt the decision to require a criminal charge in such a context affects more than just how judicial institutions review executive determinations. The claim here is merely that one dimension along with there is a contrast between (for example) use of the criminal justice system and enemy combatant designations involves the role of court review.

⁷⁷ 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 future cases with similar characteristics. See *supra* note 7 for examples of briefs making this argument.

⁷⁸ See *Cole*, *supra* note 1, at 2567.

⁷⁹ The reference to “traditional judicial review” implies that executive decisions can be reviewed through judicial fora, where judges tend to believe that court decisions governing how stringent is the review of discretionary executive decisions will affect every future case (or nearly so) in a particular class, and where litigants seek, and the court can deliver, some kind of relief, such as vacating a particular government action or providing an injunctive remedy.

the media sometimes have reason to generate that information through investigations that may turn up some potentially useful tidbit of executive malfeasance (from their perspective). There is no compelling reason to think that these erstwhile investigators will routinely face sufficient incentives to investigate, which is why observers assign such relative importance to judicial review as a means of generating information with the potential to provoke political and journalistic responses.⁸⁰ Even the otherwise laudable Freedom of Information Act (FOIA) is no reliable bulwark against bureaucratic misconduct. Such misconduct may afflict responses to FOIA requests in ways that are difficult for courts to monitor, and the law itself is riddled with exceptions.⁸¹ How else might legislators and the public undertake such supervision?

i. Envisioning the Institutional Design of Audits of 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, it may be possible to glean a better sense of the practical and political opportunities to nudge bureaucracies with discretionary legal powers away from failure. Audits are not the only means of generating information about the exercise of discretion.⁸² Nonetheless, as will become clear, the exercise of designing an audit system also serves to clarify the consequentialist goals that are supposed to be served when discretion is reviewed.

As the term is used here, an audit of executive discretion is a sustained, careful evaluation of a discrete decision drawn from a larger pool. Its aim is to uncover, for each reviewed case, whether a particular discretionary decision is in accord with some defensible standard grounded in public representations of the executive branch, implicit in statutes or constitutional doctrine, or defined by the auditor in advance. In contrast to financial or more wide-ranging management audits, the audits of targeted executive 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 executive discretion would evaluate the information supporting the decision, its origins and reliability, contradictory information, and the broader context in which the decision took place. Even in instances where the population of cases from which a sample could be drawn is relatively small, audits have the potential to “increase the information extracted from [an organization’s] own limited historical experience, by treating unique historical incidents as detailed stories rather than single data points.”⁸³ Though existing 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.

⁸⁰ See *id.* (discussing the role of judicial review in generating information that can galvanize political and journalistic responses to problematic executive branch discretionary decisions).

⁸¹ See, e.g., Christina E. Wells, “National Security” *Information and the Freedom of Information Act*, 56 ADMIN. L. REV. 1195 (2004); Amy E. Rees, Note, *Recent Developments Regarding the Freedom of Information Act: A “Prologue to a Farce or a Tragedy or, Perhaps Both*, 44 DUKE L.J. 1183 (1995).

⁸² Robust ombuds systems, such as those used in Scandinavian countries, may represent another alternative. Nonetheless, if an ombuds system functioned as it is traditionally understood in being driven by public complaints, it would have some of the same strengths and weaknesses that the litigation process would, and would thereby provide a somewhat distorted picture of bureaucratic activity.

⁸³ James G. March et al., *Learning From Samples of One or Fewer*, 2 ORGANIZ. SCI. 1, 2 (1991).

Audits associated with taxation are among the most familiar. They take place in some form in most reliable tax collection systems.⁸⁴ Many tax audits are not entirely random, which reduces their ability to provide a reliable picture of public behavior. 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.⁸⁵ 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.⁸⁶

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 is particularly likely to be responsible for that gap.⁸⁷

Audits also show 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 occasionally use audits to calculate the amount.⁸⁸ Courts reviewing this practice have repeatedly endorsed it, finding neither a conflict with the statute nor one with due process.⁸⁹ “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.⁹⁰

Something similar happens in the private sector, where “your call may be monitored for quality assurance.” The point of such audits is not immediately to stop abuses or mistakes in discrete cases. It is to enable greater learning about happens to the hundreds or thousands of individuals interacting with a company’s workers, and how those interactions can be improved. Insurance companies sometimes perform a process of “closed file review,” where they spend more money figuring out whether the amount

⁸⁴ See generally JOEL SLEMROD AND JON BAKIJA, TAXING OURSELVES: A CITIZEN’S GUIDE TO THE DEBATE OVER TAXES (2004).

⁸⁵ 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).

⁸⁶ Slemrod and Bakija, *supra* note 32, at 175.

⁸⁷ *Id.* (“[T]he estimates are based on data that is now over fifteen years old. But these are the best numbers around.”).

⁸⁸ See *Chaves Co. Home Health Serv., Inc. v. Sullivan*, 931 F.2d 914 (DC Cir. 1991).

⁸⁹ *Id.*

⁹⁰ 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).

of money paid out for a particular insurance claim was correctly calculated than they do paying out the claim itself.⁹¹ It is not difficult to see why managers would rather know more about how their employees are performing. 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.⁹²

Private sector employers value audits for a reason. Audits mitigate the problem, common to public agencies regulating private behavior, of learning how individuals are actually functioning in an inherently complex and unpredictable environment. Indeed, if regulators avoided random auditing techniques altogether, they would face at least two problems. Existing knowledge about where problems lie may prove deficient or outdated. Perhaps more important, strategic actors can simply evade review by avoiding domains where enforcement is already occurring. This is why the Supreme Court lauded random enforcement in *United States v. Biswell*.⁹³ One afternoon, a pawn shop owner who was federally licensed to deal in sporting firearms was surprised to find a Treasury agent arrive to inspect the premises. In holding that the Treasury agent could do so without a warrant, the majority observed that:

[I]f inspection is to be effective and serve as a credible deterrent, unannounced, even frequent, inspections are essential. In this context, the prerequisite of a warrant could easily frustrate inspection; and if the necessary flexibility as to time, scope, and frequency is to be preserved, the protections afforded by a warrant would be negligible.⁹⁴

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. Occasionally, government agencies audit the performance of their own workers.⁹⁵ 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.⁹⁶ 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.⁹⁷ Like the GAO, the Inspectors General have the legal power to investigate how federal officials use their targeted discretion. The existence of these structures

⁹¹ See J. David Cummins and Sharon Tennyson, *Controlling Automobile Insurance Costs*, 6 J. ECON. PERSP. 95 (1992). For a more general discussion of the value of, and incentives for use of, audits in the private sector, see Ross L. Watts and Jerold L. Zimmerman, *Agency Problems, Auditing, and the Theory of the Firm: Some Evidence*, 26 J. LAW & ECON. 613 (1983).

⁹² See Kreps, *supra* note 33, at 763-64.

⁹³ 406 U.S. 311 (1972).

⁹⁴ 406 U.S. at 316.

⁹⁵ See Exec. Order No. 12,564, 51 FED. REG. 32889 (Sept. 15, 1986)(subjecting numerous categories of federal employees to random drug tests); *Guilty Until Proven Innocent*, N.Y. TIMES (Jan. 20, 2006)(describing how the IRS Office of the Taxpayer Advocate audited a random sample of 500 tax returns where refunds had been frozen and taxpayers complained, and found that 66% of taxpayers deserved a full refund and another 14% deserved a partial refund).

⁹⁶ See John T. Rourke, *The GAO: An Evolving Role*, 38 PUB. ADMIN. REV. 453 (1978).

⁹⁷ See PAUL LIGHT, *MONITORING GOVERNMENT: INSPECTORS GENERAL AND THE SEARCH FOR ACCOUNTABILITY* 42 (1993).

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, discussed below.⁹⁸

To understand how audits of executive discretion would work, imagine a world much like our own, where only some decisions are subject to stringent judicial review, and others are subject to less stringent review. For fairly obvious reasons, political principals (to use the parlance of game theory) desire to know how the government is using its discretion. But constraints exist in the form of a limited budget to review decisions, and concerns 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.⁹⁹

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 occupational safety administrative compliance orders, or all decisions to prosecute or not to prosecute case referrals from law enforcement agencies regarding mail fraud). 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 full potential 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 (for instance, in the criminal context, the auditor’s determination of a standard would be shaped by statements of prosecutors regarding the purposes that prosecutions). Fourth, the results of the audit would be made available to legislators and the public, a development that could (under certain conditions) help pressure the agency to make modifications in its conduct.

⁹⁸ See *infra* Part III.

⁹⁹ The proposal is, therefore, somewhat reminiscent of one that Mashaw offered in the concluding pages of his exhaustive study of Social Security disability benefit determinations. See JERRY MASHAW, BUREAUCRATIC JUSTICE 226 (1983).

Intelligence information could be used in these determinations because it can be reviewed in camera.¹⁰⁰

Admittedly, the effect of this institutional mechanism depends on whether legislators and the public react to the audits. While both might sometimes ignore those results, the media's reaction to GAO and Inspector General reports suggests that audits could prove to be salient.¹⁰¹ Judicial review would continue in the background at whatever standard of deference courts and legislatures choose. Judges might even evaluate executive clamoring for deference by weighing whether a reliable audit system is in place, and in others courts might approach their cases differently as a result of what the audits revealed. Although audits would not necessarily provide relief to every aggrieved person or group, they would help legislators, organized interest groups, and the public to learn far more about what government does than is currently known.¹⁰²

ii. Trading Off Breadth Against Depth Assuming Marginal Cost Analysis

The potential value of audits is readily apparent if one makes the heroic assumption that elaborate analyses of marginal costs and benefits in fact determine standards of stringency governing a pool of potential cases. Assume for a moment, therefore, that legislators and courts have appropriately weighed the costs and benefits of reviewing a certain kind of decision at a particular degree of stringency. They have decided, for example, that orders freezing assets should get nothing more than highly deferential arbitrary and capricious review. Even if conscientious courts and legislators think of this degree of stringency as the best way to strike a balance among competing concerns, they may still recognize that there is an unfortunate by-product of this choice of stringency level. Specifically, some types of errors associated with these decisions can only be detected if review is more stringent than at present. When people engage in deliberate wrongs, for example, they tend to make efforts to hide their misconduct. In effect, the function mapping stringency of review to probability of detecting mistakes or manipulation can be radically discontinuous, in which case it may be better to manage the costs of review by reviewing fewer cases more thoroughly. It is precisely the trade-off social scientists make when they consider whether to allocate scarce resources to getting a larger sample or to investigate their cases more profoundly,¹⁰³ and that courts themselves occasionally make when they take cautiously use samples of claims in a class, or health care reimbursement requests to learn more in the litigation process.¹⁰⁴

Of course, audits introduce their own costs into the equation. While fixing the precise cost of an audit system depends on institutional design issues taken up in Part II,

¹⁰⁰ 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).

¹⁰¹ See *infra* Part III.a.

¹⁰² 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.

¹⁰³ See *infra* notes ___ (emphasizing the potential value, in social research design, of using scarce resources to generate more information about individual cases rather than increasing sample size).

¹⁰⁴ See *infra* notes ___ (discussing instances of sample adjudication).

there is good reason to expect that those costs would be lower than those associated with an expansion in the availability or stringency of traditional judicial review. For one, the costs of audits are “scalable,” such that we can audit a smaller proportion of cases if the costs are perceived as being too high, then the proportion of cases that are audited could be reduced. If one takes the language of judicial opinions and briefs (particularly from the government) seriously when they discuss the appropriate stringency of review. Beyond the reflexive invocation of precedent, there appears to be a consequentialist concern about how review at a particular degree of stringency will generate costs. Presumably those costs would be less – even if we kept the potential remedy – if only one of the whole universe of potential cases were audited, or two or three out of a pool of hundreds or thousands. If one takes the language from the judicial process seriously, a logical implication seems to be that reviewing all potential cases is more costly than reviewing some fraction of them. Moreover, audits would not necessarily yield a direct remedy.¹⁰⁵ To the extent that analyses of the marginal costs and benefits of review incorporate the possible costs to the executive branch of having a decision vacated by a court, then audits would also involve lower costs because they can be designed merely to reveal information rather than to provide direct relief.

iii. Alternative Assumptions: Separate Informational and Cost Aspirations

While painstaking analysis of marginal costs and benefits may amount to a compelling prescription for deciding on stringency of review, such an analysis is almost certainly a far-fetched description of how courts and legislatures actually make decisions. Instead, as policymakers weigh the consequences of discretion, they almost certainly recognize that various actors -- from government actors, to directly aggrieved parties, to the public at large – are likely to bear different burdens, and may reap quite different rewards, from the review of discretionary decisions. This complicates marginal cost analysis in a host of predictable ways. In the absence of far more precise moral or analytical guidance than currently exists, courts and legislatures may find it nearly intractable to ask (for instance) how many mistaken enemy combatant designations are worth tolerating in exchange for a given marginal increase in security that is allegedly only possible if some proportion of mistaken designations is tolerated.¹⁰⁶

Given these difficulties, one might therefore use a different metaphor to describe the work of principled courts and legislatures when they make decisions about the proper stringency of review. One might imagine that society (through its courts and legislatures) makes *two* initially separate calculations when deciding on how to review decisions. One sets an aspiration that a given review procedure (such as judicial review) not exceed a

¹⁰⁵ If determined advocates of executive branch discretion choose to treat the potential loss of public trust arising from as the most important cost to be borne, then the very fact that audits might be effective in changing public perceptions about executive discretion might allegedly still make them too costly. This argument may not always be entirely disingenuous, but it is a harder one to defend than one rooted in a host of costs involving bureaucratic attention, resources, frivolous claims, and similarly tangible costs. In any case, the specificity gained in discussing costs may be at least a partial reward from contemplating audits.

¹⁰⁶ The argument is not that such conversion is impossible, in the sense that either the costs and benefits or the relevant interpersonal utilities are incommensurable. Instead the contention is that decisionmakers, bombarded with briefs and policy arguments from parties with competing positions, probably find it simpler to set a threshold of total costs that a given review process should not exceed, and a separate threshold reflecting the average informational benefits sought from review of a given case.

maximum acceptable total cost. And another sets an aspiration regarding the minimum acceptable probability that a problem (whether an innocent mistake or a willful manipulation) will be discovered in the average reviewed case.

A simple model helps illustrate the potential differences between audits and traditional judicial review in a world where society sets such separate informational and cost aspirations. Begin by accepting the executive branch's premise that more intrusive judicial review is a problem because it inordinately raises the cost of review associated with the average case, 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 the total cost of reviewing instances where executive discretion is used. In this context, the reference to cost implicates several factors. 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-detering may also be greater as the proportion of cases eligible for review rises towards 100%. On the other hand, more stringent review is valuable because it is more likely to reveal problems in discretionary decisions.

To 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 $Pr()$ denote the probability that a problem is discovered in an average case 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. And imagine that the following plausible conditions hold:

- 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 stringency of review dimension S running horizontally.
- For any given set of cases that are reviewed with a particular degree of stringency, 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*.
- $Pr()$ increases as a linear function of S associated with some minimum number of cases reviewed, such that one can *also* define a two dimensional space consisting of a probability of discovering problems dimension $P()$, and a precisely orthogonal stringency of review dimension S .
- At some levels of stringency, the probability of discovering problems falls below the social aspiration for 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 social aspiration for the maximum cost of review that society is willing to bear for a given method of reviewing executive discretion. Because C increases as a function S , C_{max} corresponds to a maximum degree of stringency of review (call this S_{max}).

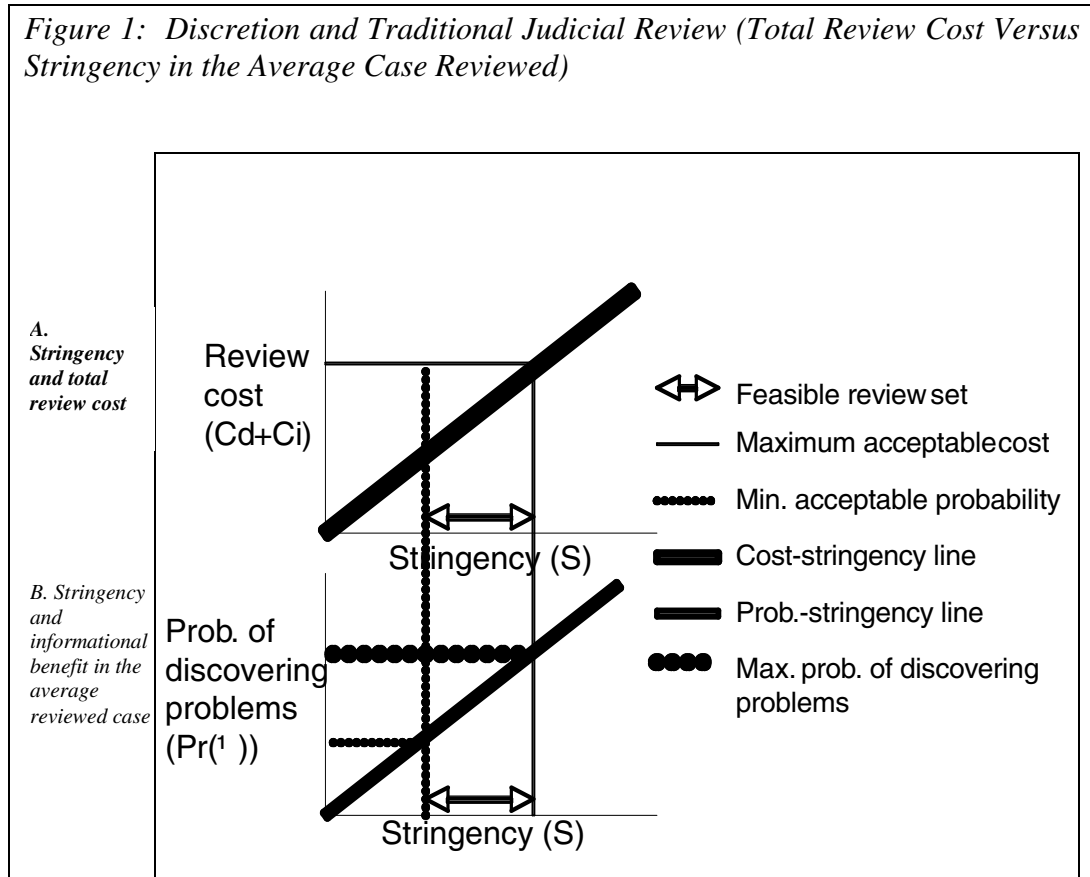
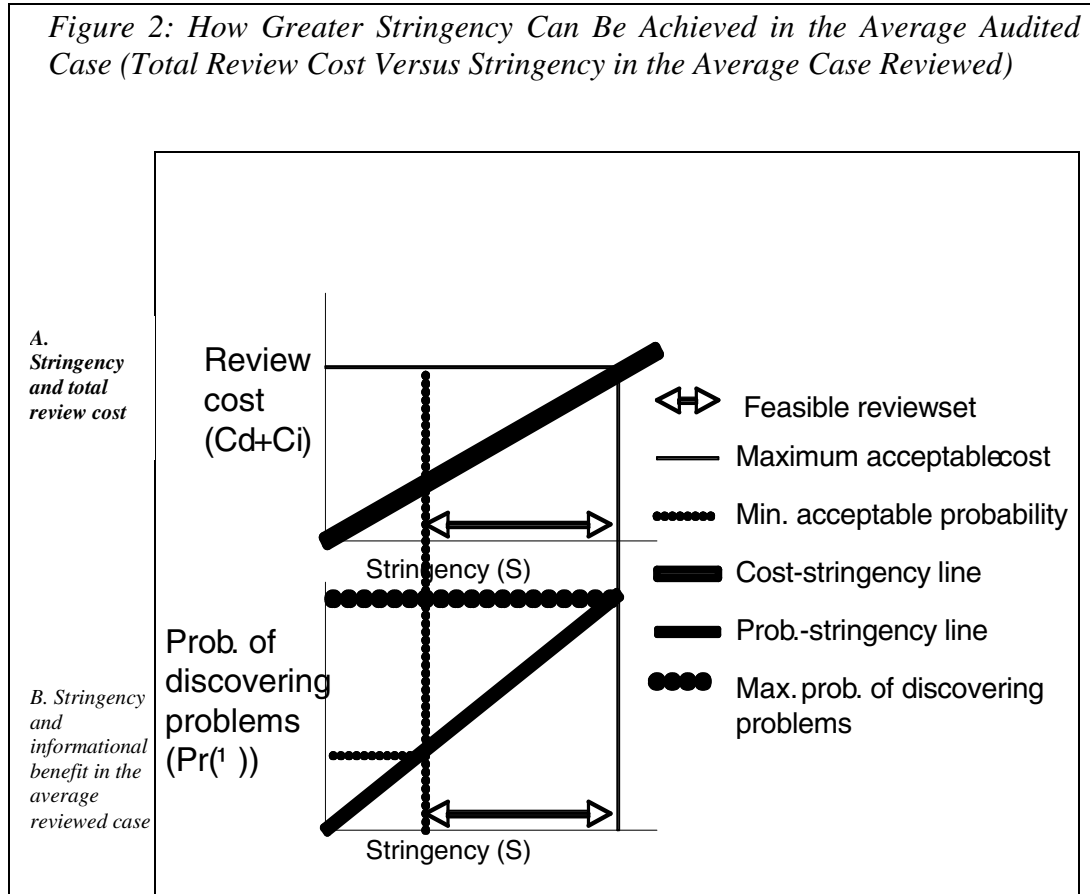


Figure 1 shows how these factors can interact in situations where separate goals have been set for managing the total cost of review and the probability of discovering problems in the average reviewed case, and judicial review is nonetheless likely to work reasonably well. Any point on the horizontal stringency axis (at the top and bottom of the figure) would have corresponding points, indicating the total cost of that degree of stringency and its associated probability of discovering problems in the average case, on the two perpendicular axes. The upper part of the figure shows the relationship between the *total* costs of review and stringency for 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

in an *average* case.¹⁰⁷ The space between the lowest acceptable bound of review (S_{min}) and the highest permissible cost (which corresponds to S_{max}) is what might be called 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 will not 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.¹⁰⁸



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 for an average case 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

¹⁰⁷ Thus, in this first figure $C=S=Pr()$.

¹⁰⁸ 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.

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.¹⁰⁹ Now 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 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) can choose a level of stringency above S_{min} – that is, above the level of stringency associated with the minimum socially-acceptable probability of discovering problems. Although this is being achieved for only a proportion of cases (in Figure 2), the information produced about that subset of cases is vastly more valuable (assuming the particular conditions) than what is learned from the essentially rubber-stamp review of every decision.¹¹⁰

The preceding discussion highlights two major differences between audits and traditional forms of judicial review. The costs of audits can be adjusted by changing the number of cases reviewed, 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 explained below, both of these characteristics turn out to be important in making the case for supplementing judicial review with audits of executive discretion.

D. Why Traditional Judicial Review Suffers From Limitations

So far the argument has proceeded by showing how audits can substitute some of the functions associated with judicial review. The properties associated with audit review, moreover, can avoid certain distortions common to the judicial process. But the value of such properties cannot be appreciated in the abstract. To better understand such 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. This entails a careful accounting of the alleged value of robust executive power in the absence of judicial scrutiny, along with an evaluation of the process that may undermine desirable bureaucratic behavior when the executive branch has been entrusted with such additional discretion. Such analyses make it easier to appreciate the strength of the case for auditing executive discretion, and the concomitant problems associated with a world devoid of audits.

i. Reviewing Rationales for Departing From a Limited-Discretion Baseline.

Given the allegedly intimate link between accountability and court review, departures from a baseline of stringent review presumably should be contingent on a satisfactory accounting of the benefits from such a move. So what are the benefits? A fairly obvious one is speed, or what is commonly termed “efficiency.” Some decisions

¹⁰⁹ In contrast with the previous figure, the relationship is now roughly $C=3/2 S = 3/2 Pr()$.

¹¹⁰ See *infra* notes ___ for a more detailed explanation of this point.

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 tropical 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.¹¹¹ 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.¹¹² On a related note, less review also saves two kinds of resources: those 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 are fashioning a standard that will apply to all (or nearly all) similar cases.¹¹³

The argument for discretion in such cases often exalts the centrality of expertise.¹¹⁴ The conventional wisdom is that agencies and the executive branch have greater specialized technical competence than the judges who might review their decisions. Given its perceived intellectual pedigree, this justification recurs in judicial decisions in a wide range of domains, regardless of whether the subject is medical evaluation of disability claims, military planning, or evaluation of chemical data.¹¹⁵ 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, or innovate in desirable ways that may not immediately inspire public confidence.¹¹⁶ Discretion may also have a role in helping government harmonize competing goals, trading off some desired goals against further delays (for example) in achieving policy objectives considered less compelling.¹¹⁷

In the same vein, supporters of executive discretion accept bold suppositions about executive branch accountability to bolster their case. Accountability is surely a contestable and often ambiguous concept. But scholarly references to it appear to encompass, at a minimum, the idea that the public should be able to assign responsibility for government decisions and to force decisionmakers to bear a cost when those decisions are not acceptable. In an ironic twist, the rhetoric of accountability that so often bolsters arguments for stringent judicial review sometimes serves precisely the opposite goal.

¹¹¹ See *Water Keeper Alliance v. U.S. Dept. of Defense*, 271 F.3d 21 (1st Cir. 2001).

¹¹² See Cornelius Kerwin, *Rulemaking* (3^d ed. 1999). See also McGarity, *supra* note 22, at 21-23.

¹¹³ See *supra* note 28.

¹¹⁴ 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).

¹¹⁵ See *Vermont Yankee Nuclear Power Corp. v. NRDC*, 439 U.S. 961 (1978); *I.N.S. v. Lopez-Mendoza*, 468 U.S. 1032 (1984).

¹¹⁶ Cf. KEITH KREHBIEL, *INFORMATION AND LEGISLATIVE ORGANIZATION* (1991).

¹¹⁷ See *Greenbaum v. E.P.A.*, 370 F.3d 527 (6th Cir. 2004).

The argument proceeds along the following lines. 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 punishing the executive (or the inferior officer) for her decisions.¹¹⁸ 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."¹¹⁹

ii. Why Bureaucracies Granted Discretion Fail (to Learn).

The arguments extolling discretion each have a grain of truth. What they do not 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. But that belief is not always fully explained, and court review carries concomitant risks of over-detering executive branch activity. 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 arise, and that judicial review helps create conditions that foster learning.

Four separate but interrelated reasons support this claim. First, a substantial body of research suggests that people learn when they have reason to do so.¹²⁰ Other things being equal, the dilution of review may deprive individuals in public bureaucracies of reasons to learn (at least, limits on review may disrupt public officials' incentives 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.¹²¹ Second, organizations develop routines that blind them. As Diane Vaughan wrote in her study of the *Challenger* launch decision:

¹¹⁸ 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).

¹¹⁹ *Newman v. United States*, 382 F.2d 479 (D.C. Cir. 1967).

¹²⁰ See ARTHUR LUPA AND MATHEW MCCUBBINS, *THE DEMOCRATIC DILEMMA: CAN CITIZENS LEARN WHAT THEY NEED TO KNOW?* (2001).

¹²¹ 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 128 and 129.

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 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.*¹²² [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 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.¹²³ 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.¹²⁴

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,¹²⁵ but the basic insight is a simple one. Executive authorities face fewer checks in the domain of discretionary action than in traditional regulatory or criminal justice realms. Accordingly, 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

¹²² DIANE VAUGHAN, *THE CHALLENGER LAUNCH DECISION: RISKY TECHNOLOGY, CULTURE, AND DEVIANCE AT NASA* 409-410 (1996)

¹²³ See Michal Tamuz, *Learning Disabilities for Regulators: The Perils of Organizational Learning in the Air Transportation Industry*, 33 *ADMIN. & SOC.* 276 (2001); March et al., *supra* note 31, at 2-3.

¹²⁴ See Mashaw and Harfst, *supra* note __, 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.

¹²⁵ See Mariano-Florentino Cuéllar, *The Mismatch Between State Power and State Capacity in Transnational Law Enforcement*, 22 *BERKELEY J. INT'L L.* 15 (2004).

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. It may not. This state of affairs may skew citizens' ability to evaluate the effectiveness of their own government. And the discretionary actions may themselves have costs, including the creation of perverse incentives for regulated groups,¹²⁶ 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 (at least for some) in freedom. There is, finally, the specter of more deliberate transgressions. 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.¹²⁷

When stalwart defenders of executive discretion come close to acknowledging these realities, their most frequent move is to invoke a political process that is rarely expressly defined. They are obviously right to recognize how public organizations exist in a larger political context. But assuming that such a context will reliably and consistently counterbalance the tendencies I've just described requires accepting heroic assumptions. Even if voters often behaved relatively rationally as the term is conventionally understood in modern political science, the results of the political game are endogenous to the information available. Agency relationships change in response to what the players come to know, even if – to paraphrase Defense Secretary Donald Rumsfeld – voters know what they do not know.¹²⁸

A more plausible assumption can be grounded in the extensive behavioral research tradition in mid-to-late 20th century political science, suggesting that voters often do not know what they do not know.¹²⁹ One might even question the electorate's distribution of its scarce cognitive attention, as there is no particularly good reason to think that voters come to focus on the facets of law or policy that they should even if we use their own consistently expressed and stable values as a benchmark.¹³⁰ These limitations constrain the electorate's capacity to provide a bulwark against bureaucratic failure. And they explain, among other things, why legislators themselves often do just fine not only if they ignore festering problems of bureaucratic competence,¹³¹ but if they

¹²⁶ *See id.*

¹²⁷ *See* THE ABU GHARIB INVESTIGATIONS: THE OFFICIAL REPORT OF THE INDEPENDENT PANEL AND PENTAGON ON THE SHOCKING PRISONER ABUSE IN IRAQ (2005).

¹²⁸ *See* Hart Seely, *The Poetry of D.H. Rumsfeld*, SLATE (April 2, 2003), avail at. <http://www.slate.com/id/2081042/> (last accessed December 13, 2005).

¹²⁹ *See*, e.g., ARTHUR LUPIA & MATHEW MCCUBBINS, *THE DEMOCRATIC DILEMMA: CAN CITIZENS LEARN WHAT THEY NEED TO KNOW?* (1998)(emphasizing citizens' limited knowledge about politics); Richard R. Lau and David P. Redlawsk, *Voting Correctly*, 91 AM. POL. SCI. REV. 585 (1997) (using survey data to construct a statistical model estimating the proportion of voters that vote "correctly" given their attitudes and impressions, and finding that one in four voters voted incorrectly in the five American presidential elections between 1972 and 1988).

¹³⁰ *Cf.* Mariano-Florentino Cuéllar, *Rethinking Regulatory Democracy*, 57 ADMIN. L. REV. 411 (2005).

¹³¹ *See* Kenneth E. Scott and Barry R. Weingast, *Banking Reform: Economic Propellants, Political Impediments*, in *REFORMING FINANCIAL INSTITUTIONS AND MARKETS IN THE UNITED STATES* 19-36 (George G. Kaufman, ed., 1994).

deliberately create them.¹³² The point is not that electoral checks are irrelevant. It's that they're not entirely reliable, and they are often dependent on some mechanism – like judicial review – to focus public attention and produce information for it.

E. How Audits Can Ameliorate the Limitations of Traditional Judicial Review.

The preceding discussion implies that bureaucracies should frequently be expected to face pressures to render poor decisions, unless they are subjected to constraints that judge them in promising directions. Those constraints depend, in turn, on the public availability of information about bureaucratic performance. Once the executive discretion problem is thus recast, it becomes plain how audits can often generate such valuable information. If the costs of review only apply to a small fraction of cases, then marginal cost analysis permits more stringent review of that random subset of cases. Willful malfeasance will be harder to conceal, and (more generally) mistakes that would simply not appear under deferential judicial review may emerge, regardless of whether they involve compliance orders, passenger prescreening procedures, or other forms of administrative action.¹³³

But an even more striking rationale for audits emerges if we imagine that courts and legislators are not precisely balancing marginal costs and benefits when deciding how to review executive discretion. Return instead to the possibility that legislators and courts are instead trying to negotiate the difficult terrain created by inconsistent social aspirations. What happens if a reasonable amount of review is simultaneously prohibitively expensive and extremely valuable? Lawyers may seem to be talking past each other in such a scenario, and judicial review itself may be condemned to fail because the costs of review make the socially desired degree of stringency for reviewed cases appear too expensive. Such costs may arise from the resource burdens directly associated with the provision of review, from increases in the probability of interference with desirable executive branch activity, or from the resources necessary for the executive to respond to judicial proceedings. Figure 3 indicates the problem. It depicts a situation where the curve depicting the relationship of stringency of review to costs of review (i.e., the cost-stringency function) actually looks quite different from the curve depicting the relationship between stringency of review and the probability of uncovering mistakes (i.e., probability-stringency function).

Suppose, for example, that the first function is a linear one, just as in Figures 1 and 2. Stringency pushes up the cost of review in a constant fashion. More stringent review means more time spent adjudicating disputes about executive discretion, more executive branch resources spent resolving those disputes, and a greater probability that

¹³² See Barry R. Weingast, *Caught in the Middle: The President, Congress, and the Political-Bureaucratic System*, in *THE EXECUTIVE BRANCH* 312 (Joel D. Aberbach and Mark A. Peterson, eds. 2005).

¹³³ This assumes, plausibly, that for most domains the gains from greater stringency in review of the limited sample of cases outweighs cursory review of all cases in a particular class (which is already available, in most cases, with existing deferential forms of judicial review). Even if judicial review is entirely precluded, stringent review of a small sample may prove vastly more informative than dividing a scarce review budget among an entire pool of potential cases. For instance, a non-linear function mapping stringency to a given probability of discovering problems may doom deferential review to uncover nothing; meanwhile, a small percentage sample of a larger population may reveal almost as much as review of the entire population. See *infra* Part II.b for an elaborate of the last point.

review will render decisions about freezing assets or administrative compliance orders unworkable. 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 $Pr()$, 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 (such as TSA officials trying to cover up problems with how they put names on a “no-fly” list). It might also describe instances where bureaucratic requirements appear to have been followed on paper, but where, in fact, those requirements are not actually followed.¹³⁴ In such cases, it is likely that problems could only be detected with a decidedly more exacting brand of review.

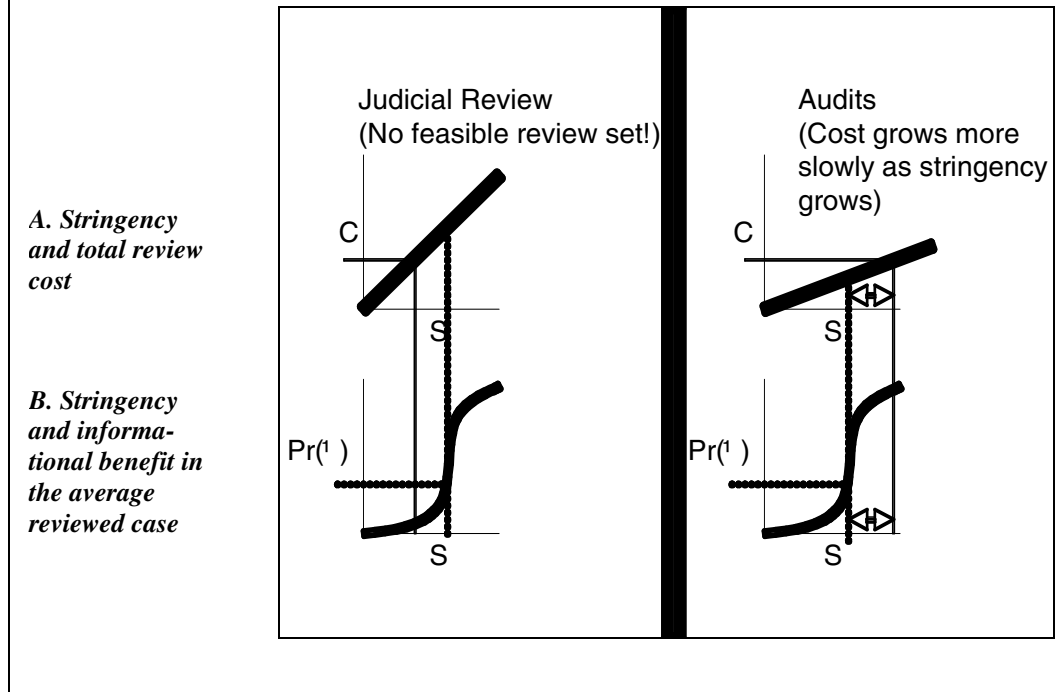
Now imagine juxtaposing this situation with a social aspiration for a relatively low total cost (S_{max}), which is what the court may set if the decision has to do with national security. And notice what this does to the feasible review set. It disappears, leaving the reviewing court with no way to simultaneously achieve social aspirations regarding the maximum total cost of review and the minimum acceptable probability of discovering problems for a reviewed case. 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 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 number of decisions reviewed is also smaller. But in many domains, it seems plausible that greater stringency for a smaller number of decisions would tend to be associated with a higher comparative probability of detecting problems than lower stringency for a larger number of decisions. Such a trade-off might be especially worth making, for example, when the total number of decisions rendered is relatively smaller than in other domains (thereby making a small number of audited decisions into a larger proportion of the total decisions), or when the soundness of a decision depends not on what an administrative record says but on how the record was gathered. In situations where the stringency increase possible from reviewing fewer cases yields such a greater probability of uncovering problems, the result of using audits creates the possibility of a small feasible review set where there was none before –

¹³⁴ After the September 11 attacks, for example, federal law enforcement authorities detained hundreds of aliens in the United States and established a procedure to prevent a detainee’s release until the FBI determined that a given individual was “cleared” of involvement. Despite the authorities’ initial protestation that individuals were released promptly after receiving FBI clearance, it actually took between 5 and 119 for individuals to be released. The 119 day hold was the result of an “administrative error.” See U.S. DEPARTMENT OF JUSTICE, OFFICE OF THE INSPECTOR GENERAL, THE SEPTEMBER 11 DETAINEES: A REVIEW OF THE TREATMENT OF ALIENS HELD ON IMMIGRATION CHARGES IN CONNECTION WITH THE INVESTIGATION OF THE SEPTEMBER 11 ATTACKS 129 n.109 (April 2003), avail. at <http://www.fas.org/irp/agency/doj/oig/detainees.pdf> (last accessed, November 1, 2005).

enough room for the auditor to avoid exceeding the maximum cost yet to still have a reasonable opportunity to detect problems afflicting discretionary decisions.

Figure 3: Instances Where Judicial Review “Fails” in the Average Case (Where Society Generates Separate Informational and Cost Aspirations)



The value of that error detection mechanism would also lie in the more representative picture of bureaucratic behavior it might yield. 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.¹³⁵ As Bill Simon has pointed out in the welfare context, for example, this pattern can gradually distort public perceptions of a policy’s strengths and weaknesses by focusing attention on only one aspect of an administrative system’s problems (e.g., underpayments, rather than overpayments).¹³⁶ In contrast, audits have the potential to generate information regardless of the willingness or opportunity of individuals to challenge specific decisions.

¹³⁵ 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).

¹³⁶ See William H. Simon, *Rights and Redistribution in the Welfare System*, 38 STAN. L. REV. 1431 (1986).

While the preceding features are likely to render audits quite valuable, two caveats are in order at this point. First, the argument thus far tenders audits as a conceptual alternative to judicial review, not as a wholesale replacement of it. Its principal contention is that we should consider the implications of policing discretion through audits instead of through judicial review, which I've characterized as a form of supervision premised on applying the same degree of stringency to every case in a given class. This does not imply that audits should generally replace traditional judicial review as an exclusive means of overseeing executive discretion. Judicial review obviously serves a host of important functions, of which producing information about the performance and reasonableness of public bureaucracies making discretionary decisions is only one. Litigation harnesses the intricate machinery of adjudicatory bureaucracy to articulate and clarify legal norms in the context of specific cases. It 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.¹³⁷

Second, the belief that audits can deliver the aforementioned benefits depends on making certain assumptions about the powers of the auditor, though none are particularly 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. In this context, the reference to "better position" implies at least three qualities: (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.¹³⁸ Nonetheless, the auditor(s) must know

¹³⁷ There may be some instances where actual or perceived resource constraints force a choice between policing discretion through audits or doing so through some expanded version of judicial review (i.e., supplementing narrow review of constitutional questions with broader arbitrariness review). In those exceptional circumstances, the framework I have provided suggests that sometimes – such as when it's likely that officials have made an effort to conceal willful malfeasance – allocating scarce resources to audits would be preferable to allocating them to traditional judicial review. Nonetheless, my primary goal is therefore to argue that supplementing such review with audits would make possible the production of socially valuable information. Such information simply would not be available where reviewing authorities are constrained to review every case in a particular class with the same degree of stringency. Once that information is available, judges could be among the consumers of it by adjusting their evaluations of the appropriate degree of judicial scrutiny that a given class of decisions warrants.

¹³⁸ See, e.g., Vaughan, *supra* note 64, 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).

something about what constitutes accuracy when allegedly terrorist assets are frozen or when agencies use administrative compliance orders.¹³⁹ Whether it is possible to generate this and other conditions depends in large measure on how to resolve questions about the details of the institutional design addressed below.

II.

WHY THE INSTITUTIONAL DESIGN PROBLEMS INHERENT IN AUDITS ARE MANAGEABLE

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

As Kenneth C. Davis recognized a generation ago, government authorities are flush with power to make highly informal decisions affecting people, where “the usual quality of justice” may be quite low.¹⁴⁰ Such discretion, showcased in the preceding discussion, may often involve specific, targeted decisions whose primary effect is on specific individuals and groups.¹⁴¹ This type of discretionary power government officials possess to freeze allegedly terrorist assets, place someone on a government-run “no-fly” list, or levy certain environmental or occupational safety fines could be called “targeted discretion.” 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.¹⁴² Because of their frequency and their impact on discrete individuals and groups, targeted discretionary decisions are most immediately suitable for audits.

Targeted executive 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 be carried out through the exercise of targeted discretionary decisions. 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

¹³⁹ 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.”

¹⁴⁰ Davis, *supra* note 1, at 216.

¹⁴¹ 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). Obtaining review of such decisions under the arbitrary and capricious standard, or on the basis of some constitutional theory, is unlikely to provide adequate oversight for these decisions because of the specter of the review costs problem. But one should nonetheless beware the term informal adjudication as a description of the full scope of discretionary decisions that merit new review mechanisms. Many discretionary decisions are neither informal in the sense that they are bereft of some alleged procedural safeguards (for example, review of administrative compliance orders), and in the case of many others, the term adjudication is but a euphemistic concession to an almost entirely unrealized aspiration.

¹⁴² See Brief for Respondent, *Hamdi v. Rumsfeld*, *supra* note 7, at 12.

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 given 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 – should not be made entirely public.¹⁴³ They might also argue that, in some domains such as presidential decisions involving executive agreements, standards reflecting policy judgments should develop organically in response to experience instead of being fixed ahead of time, or that any review system at all would wreak havoc.¹⁴⁴ 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 the following classes of decisions: decisions that explicitly disavow consistency with *any* standard, decisions that fix a standard that is supposed to apply across cases, and decisions that *apply* standards (or even quite general values) to specific cases. To the extent that executive authorities claim to be applying a standard, then arguments against review become exceedingly difficult to accept. Though some may even insist that certain discretionary decisions involving national security (for example) are entirely immune from any standard,¹⁴⁵ in most cases such a claim seems hard to reconcile with a simple but persistent imperative: that government decisions should not be arbitrary. Indeed, audits of executive discretion may prove viable even when the decisions in question superficially appear less amenable to sampling. 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, by aggregating cases from different domains into a larger population from which to sample.¹⁴⁶

¹⁴³ See *Doherty v. U.S. Dept. of Justice*, 775 F.2d 49 (2d Cir. 1985). For a skeptical view of such contentions, see Vaughan, *supra* note ____.

¹⁴⁴ *Cf. Marbury*, 5 U.S. at 170.

¹⁴⁵ 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.

¹⁴⁶ Small samples are not a fatal problem when Bayesian techniques are applied and the analysis is accompanied by appropriate assumptions. See, e.g., 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).

B. Most Institutional Design Problems Have Plausible Solutions

While a framework for conducting audits is adaptable almost by definition, practical questions of institutional design lurk just beneath the surface. The most immediate such question concerns what discretionary decisions should be prioritized for auditing in a world of scarce resources. It is plainly obtuse to seek audits of the most trivial discretionary actions and informal adjudications.¹⁴⁷ Even if this were not so, scarce resources should compel auditors to set rough priorities. In doing so, auditors should probably draw on at least four factors reflecting defensible intuitions about the underlying goals of the audit system: (1) the extent (or absence) of alternative review mechanisms, or factors favoring interest group or legislative reactions likely to reduce the probability of bureaucratic failure, (2) the costs of discretionary decisions (particularly erroneous ones) to individuals or groups aggrieved, (3) the existence of a standard against which decisions could be assessed, or the relative feasibility of constructing such a standard from sources such as public executive branch declarations about the particular objectives of discretionary action in a given context (more on this below); and (4) the potential availability of information that an auditing authority could use to evaluate decisions. Such a framework avoids placing traffic stops at national parks at the same level of priority as enemy combatant designations or social security claims. It also avoids facile reliance on pre-existing, and potentially problematic, notions of where the greatest problems lie in the exercise of executive discretion.

Regardless of what discretionary decisions are audited, the question of how large a sample to use depends, as a prescriptive matter, on the costs and benefits of larger (as opposed to smaller) samples. As a practical matter, the auditor's approach to sampling might also be informed by political circumstances that translate into resource constraints (akin to the aforementioned "maximum acceptable cost" of review), and by prior beliefs about the desired deterrent effect on decisionmakers. Ordinary statistical theory offers a straightforward framework to calculate the appropriate sample given certain desired parameters reflecting the decisionmaker's desired level of confidence in the results.¹⁴⁸ Although this formulaic analytical approach does not capture all the nuances with which an auditor must contend, it does reveal some important considerations. For one, there is no textbook answer to the question of how large a sample should be. Decisionmakers do not (and ought not to) choose desired levels of confidence in a vacuum. As is true of courts and legislators structuring the analogous judicial domain, decisionmakers have reason to consider the costs as well as the benefits – which means that smaller samples may sometimes be appropriate even if larger ones are more representative. Since resource constraints can force a trade off between stringency of review (for any given case in the sample) and breadth of review, then decisions about sample size should depend, in large measure, on prior guesses about the slope and shape of the curve depicting the impact of stringency on the probability of discovering problems in decisions.

¹⁴⁷ Cf. Krotoszynski, *supra* note __, at 1069 ("Taken to its logical extreme, a decision to turn on the lights or lower an agency's office building temperature two degrees constitutes 'informal adjudication.'").

¹⁴⁸ See JUN SHAO, HANSHENG WANG, AND SHEIN-CHUNG CHOW, SAMPLE SIZE CALCULATION IN CLINICAL RESEARCH 28-29 (2003).

Those choices regarding sample size should also reflect the fact that sample size has diminishing marginal returns. Counter-intuitively, as the size of the total population of decisions increases beyond a certain point, then the proportion of that population that must be reviewed to gain a reasonably accurate picture of the whole population actually declines. Sample sizes of between five hundred and a thousand observations may provide a revealing statistical snapshot even when drawn from exceedingly large populations.¹⁴⁹ This suggests that auditing only a tiny fraction of a large population and reviewing the sample carefully might yield valuable new information. Smaller populations (for example, all the cases of the 500 or so individuals currently held in Guantanamo) pose more of a challenge, since even a sample that constitutes a higher total percentage of the population can prove less useful in making statistical inferences if the sample is numerically small in absolute terms.¹⁵⁰

Now juxtapose the insights of statistical and organization theory, and several implications emerge. If the total population of cases is large enough to allow the auditor to choose between 500 and 1000 cases without exceeding the maximum acceptable cost, then the resulting analysis will likely exhibit desirable properties of reliability even if the sample is a tiny proportion of the total population of cases (for example, the total number of indictments issued over several years, or the total number of disability determinations).¹⁵¹ If the cost of obtaining a sample of that size is allegedly prohibitive, or if the total population is too small to audit hundreds of cases, then the auditing process is best understood not as a means of obtaining a statistically reliable picture of conditions in the population, but as a kind of pilot study informing decisions about the merits of existing review procedures. And even small samples subject to audits can expand knowledge considerably.¹⁵² Finally, regardless of the size of the population, the

¹⁴⁹ For example, if a sample of 1,000 cases is drawn randomly from an infinitely large population where one half of the population hold a characteristic, the researcher has a 95% chance (or better) of obtaining a result that is plus or minus 3.1 percentage points of the actual distribution in a population. If a sample of 5,000 is drawn under the same conditions, the 95% confidence interval would be plus or minus 1.4 percentage points. And if a sample of 10,000 is drawn, the 95% confidence interval would be plus or minus 1 percentage point. See ALAN D. MONROE, *ESSENTIALS OF POLITICAL RESEARCH* 69 (2000). While increasing the sample size has diminishing marginal returns in large populations, things are more complicated with smaller sample sizes. See CAROLE SUTTON AND MATTHEW DAVID, *SOCIAL RESEARCH* 154 (2004).

¹⁵⁰ An exhaustive analysis of 25 cases of Guantanamo detainees, for example, constitutes about 5% of the total population. But in orthodox statistical terms it yields far less reliable information than a sample of 1% of the total population of 100,000. Cf. Monroe, *supra* note 114, at 69.

¹⁵¹ Even in such situations, the concept of statistical confidence ought to be treated as a species of metaphor, and applied with some nuance, since the auditor's role here is not to estimate a precise quantitative parameter but to conduct a reasoned analysis of the decision in accordance with a defensible standard. For an interesting effort to address the problem of tailoring qualitative concepts of reliability in the analogous context of property value appraisals, see Nathan Berg, *A Simple Bayesian Procedure for Sample-Size Selection in an Audit of Property-Value Appraisals*, POLITICAL ECONOMY WORKING PAPER NO. 06/04, UNIV. OF TEXAS, DALLAS (May 2004)(on file with author).

¹⁵² Cf. Sutton and David, *supra* note 114, at 154 (even a sample as small as 30 can yield useful statistical information). 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). Cf. March et al., *supra* note 31, at 2 (noting that sometimes “[o]rganizations attempt such pooling” of cases to create a larger sample from which to assess performance).

proportion of cases audited can be adjusted to achieve the related but distinct goal of deterring malfeasance by placing decisionmakers on notice that some proportion of their choices will be scrutinized.

Sample size also depends on whether audits are meant to address both over- and under-enforcement. Audits could, for example, 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* designated as a global terrorist organization, 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 problems that demand attention. First is deciding whether the expanded population of cases should include the whole universe (i.e., every charity, or perhaps 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 prove unworkable, at least initially. In response, auditors could design pilot studies, some of which could be targeted to domains where under-enforcement appears likely to be a more pronounced problem. 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 would arise if auditors further expand their mandate to include discretionary decisions not to act. Part IV returns to the question of how auditors could mitigate more general problems of political resistance over time, and explains how such resistance can have the laudatory effect of giving auditors reason to cultivate reputations for impartiality.

The prospect of overcoming political resistance to audits depends in part on what precise institutional actor shoulders the burden of auditing. To some observers, it may seem as though courts 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.¹⁵³ 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 are viewed as unable to require audits given constraints in their ability to impose procedures not explicitly grounded in statutes,¹⁵⁴ 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

¹⁵³ 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.

¹⁵⁴ *Id.*

(though, as I note below, they have largely avoided doing so). In short, while audits could be performed by existing federal audit bureaucracies, a combination of judicial and legislative innovation could lower barriers preventing Article III courts from more easily encouraging audits.

Whatever the precise organizational structure, the auditor must be invested with the power to compel production of evidence and testimony. In the absence of such power, 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, an approach that would further weaken the argument that review should be precluded because the information involved is too sensitive.¹⁵⁵ Because this problem has been so often 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.¹⁵⁶ In addition to courts reviewing the information in camera, high profile commissions like the September 11 Commission and expert working groups routinely get security clearances and access to classified information.¹⁵⁷ 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.¹⁵⁸ The auditor

¹⁵⁵ Under existing law, 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 avoided 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). Even in the prosecution of alleged Al Qaeda terrorist Zacharias Moussaoui, the federal government provided Moussaoui's defense team with the opportunity to review classified information in the context of the criminal discovery process. See A. John Radsan, *The Moussaoui Case: The Mess From Minnesota*, 31 WM. MITCHELL L. REV. 1417, 1433 (2005).

¹⁵⁶ See, e.g., *Jifry v. FAA*, 370 F.3d 1174 (DC Cir. 2004).

¹⁵⁷ See REPORT OF THE NATIONAL COMMISSION ON TERRORIST ATTACKS UPON THE UNITED STATES (2004).

¹⁵⁸ Deriving such implicit standards is exactly what the Supreme Court has discouraged in cases such as *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

could even use statements from the executive branch itself to see whether the audited cases seem to be consistent with those statements.¹⁵⁹ 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 concerns the consequence of audit results. As an initial thought experiment, imagine that the auditor labors under a default presumption that the results of its investigation will simply be announced to the public. A striking feature of audits may ultimately prove their capacity to enhance how executive authority is policed without directly delivering relief. What this would accomplish depends on the reactions of legislators and the mass public, which can vary depending on the circumstances.¹⁶⁰ Not everyone among the public would care enough about how laws are applied on their behalf to respond with indignation to audit results revealing arbitrariness in the process to determine who is allegedly raising money for terrorists or engaging in environmental violations. Nonetheless, as Part III notes, reports from existing audit bureaucracies turn out to already generate dozens of stories in national newspapers and television networks. Thus, public disclosure may have the potential to impact crucial features of the political game as the executive branch seeks to demonstrate its competence.

But audits’ touchstone is their flexibility. The auditor might be empowered to 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, and could ultimately affect their political feasibility.¹⁶¹ A third approach is for the results of audits to trigger additional procedural standards, such as review of more decisions through stratified sampling targeting areas where problems have been newly discovered. Perhaps more important, auditors and courts may gradually come to view themselves in a symbiotic relationship, as courts adjust their calculus of deference to executive action in response to audit results or the existence of audit programs for

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.

¹⁵⁹ 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).

¹⁶⁰ 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).

¹⁶¹ 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. Truth Commissions, for example, reflect an equilibrium where supporters have likely traded off explicit punitive power in exchange for political and economic resources.

particular sectors of activity, and auditors adjust the standards they use to audit cases in response to judicial elaborations of statutes or constitutional doctrines.

In any event, the choice among these alternatives is not one that can be taken in a vacuum. It is likely to depend heavily on some of the political considerations I discuss in Part IV. 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 such problems have been solved already.

III.

WHY REFORMS WOULD CONTRAST SHARPLY WITH THE STATUS QUO

The analysis began (for expositional clarity) by focusing primarily on court review of executive discretion. In reality, though courts have distinctive legal powers, they are but one component of a larger web of institutions potentially capable of overseeing executive discretion. The value of supplementing judicial review with a new program for auditing executive discretion depends largely on whether or not such audits are already commonplace. 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.¹⁶² 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 generally "no." These agencies undertake a tremendous amount of interesting and often 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.¹⁶³ The Inspector General offices in federal departments have a similarly broad mandate.¹⁶⁴ Yet they appear to rarely perform audits of executive discretion involving random (or stratified) sampling of legally consequential discretionary decisions, assessed against a defensible standard (either a pre-existing one or articulated by the auditors).

A small literature addresses the historical origin, legal jurisdiction, organization, and culture of these audit bureaucracies.¹⁶⁵ 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

¹⁶² See *infra* notes ___ - ___.

¹⁶³ 31 U.S.C. § 712 (1).

¹⁶⁴ 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.).

¹⁶⁵ See, e.g., Light, *supra* note 43; ROGER R. TRASK, DEFENDER OF THE PUBLIC INTEREST: THE GENERAL ACCOUNTING OFFICE, 1921-1966 (1996).

implement, and whether any of this gets media attention. These questions are relevant to the present project 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 could enhance their supervision of executive discretion.

Some preliminary answers emerge from the initial results of a more extensive, ongoing study of the audit bureaucracies currently underway. The results reflect an analysis of 400 Inspector General and GAO reports issued during the last five years. The reports analyzed assess the work of five major government agencies with a broad spectrum of responsibilities.¹⁶⁶ The analysis reveals audits of executive discretion (i.e., analyzing (1) a random sample (2) of executive branch legal determinations and (3) identifying or defining a standard against which to assess such decisions) to occur in fewer than 2% of the reports in the sample.¹⁶⁷ Fewer than one in five reports in the sample appear to use any sampling from a larger population, and the vast majority of these focused on the traditional financial accounting functions that convey almost no information about how an agency uses its legal discretion to affect directly the fortunes of individuals or groups.¹⁶⁸ Not a single one focuses the inquiry on domains where an agency exercises significant legal discretion, defines a standard in advance, and reviews a subset of cases at random. Some reports occasionally chronicle problems in administrative systems like those governing aviation security. Nonetheless, audits of

¹⁶⁶ The analysis employs a stratified sampling design focusing on these agencies because some of the goals of the larger empirical project depend on closely studying the interactions between the audit bureaucracies and a small number of specific agencies with differing functions.

¹⁶⁷ Four trained research assistants coded the GAO and IG reports. The stratified random sample was drawn from among reports completed between March 1999 and March 2005, and focusing on the following five departments (or their components): the Department of Homeland Security, the Department of Justice, the Environmental Protection Agency, the Labor Department, and the Department of Defense. Of the sample of 400 reports analyzed, about 68 (or 18%) used some form of random sampling. Most of these did so in the context of a traditional financial audit. Only seven reports in the total sample included audits of executive discretion. These covered matters such as the proportion of background checks conducted by the Transportation Security Agency on prospective airport screeners that met statutory standards, or the extent of FEMA payouts for certain disaster relief programs. A separate comparison sample of 50 reports randomly selected from among those dealing with departments not in the stratified sample yielded similar results.

¹⁶⁸ For a rare example of an Inspector General report that comes close to achieving the goal of systematically analyzing executive discretion, see U.S. Dep't of Justice OIG, *Detainee Report*, *supra* note 101. Although this report focuses primarily on describing the treatment of immigration detainees at two government facilities, it did undertake a thorough review of the files describing the status and treatment of the 762 aliens held in connection with the September 11 investigation. While the report does not leverage the insight that fewer cases reviewed more thoroughly can sometimes yield greater information, its searing analysis provides so informative in part (in all likelihood) because the investigators conceived of their task in terms of generating information, rather than in terms of resolving particular cases. The vast majority of other reports surveyed involve either the use of interviews and non-random analysis of documents to make broad policy recommendations, or traditional financial audits. Of course, in some sense the GAO and Inspector General reports that monitor any aspect of government performance are examining the application of executive discretion because they are examining how agencies perform in some (ordinarily quite general) aspect of a mission they are lawfully empowered to undertake. While all of these studies expand the scope of information (and in some cases generate considerable media attention), there is a striking gap in the methodology and subject matter coverage of these reports substantially overlapping with the audits of executive discretion recommended here. Specifically, what the reports almost never encompass are systematic audits of repetitive, discrete applications of executive power affecting individuals or groups using sampling.

executive discretion are essentially missing from the picture of the audit bureaucracies' work. 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.¹⁶⁹

Notice that the distinction between audits of executive discretion and most of the existing work done by the audit bureaucracies is not just a matter of *what* government conduct is analyzed, but *how* the analysis is undertaken. Procurement investigations, for example, are a predictable staple of GAO investigations given legislators' interest in dramatizing fraud and the potential cost recoveries that can help the agency justify its own budget. Most such investigations focus on areas where the agency or its legislative masters already perceive a potential problem. No examples existed among the reports analyzed where the agency picks a random sample of procurement transactions to examine how they have been carried out. The audit bureaucracies may have certain valuable pre-existing information about where problems exist. But as the tax analogy demonstrates, that information is unlikely to be pervasively correct, and in any case the selection of a random sample of transactions to audit further burdens participants trying to evade review by anticipating the agencies' priorities.

The preliminary results from the GAO/IG project also permit a partial response to two concerns likely to be raised by skeptics of audits. First, would anyone in the public actually *care* about the results of the audits? 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. An analysis of 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 reveals that the audit bureaucracies receive considerable media attention. Nearly a thousand articles during this period appearing in the New York Times mention the GAO or the IG offices. A random 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.¹⁷⁰ These figures exceed the number of mentions garnered by many federal cabinet agencies.

Second, would the agencies merely ignore the prescriptive implications of audits? If they did, then the promised learning bonus from audits would be unlikely to materialize. Although data are not yet available regarding the recommendations of the IG

¹⁶⁹ See GAO Report, *supra* note 83. For a study of Inspectors General revealing that only about 10% of their audits sought to measure program results as opposed to financial, procurement, or management performance, see Kathryn E. Newcomer, *Opportunities and Incentives for Improving Program Quality: Auditing and Evaluation*, 54 PUB. ADMIN. REV. 147, 152 (1994).

¹⁷⁰ 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.

offices, I have already gathered data on all the approximately 10,000 discrete recommendations made by the GAO over the last 15 years. The data cover recommendations to the full range of federal government agencies, from Interior to the State Department. An analysis of such data proves revealing. After the GAO makes recommendations to an agency in its reports, its staff generally conduct follow up interviews, additional investigation, document reviews, and issue queries to the agency leadership. The audit agency then determines (on the basis of these qualitative methods) whether a given recommendation is implemented sometime during the next four years.¹⁷¹ Nonetheless, about 79% of recommendations are implemented, perhaps in part because of the potential media attention reports generate. These results must be interpreted with a measure of sagacity. 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”). The adoption of recommendations is likely to be influenced by political factors, such as the extent of division in appropriations and authorizing subcommittees that oversee the agency in question. It is likely, too, 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 reject the relevance of the audit bureaucracies, even if they do not currently perform the sorts of audits that would generate critical missing information about the use and abuse of executive discretion.

B. Neither Do Legislators

Another possible setting where audits of executive discretion 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.

Legislators depend on oversight of the bureaucracy to achieve their goals. Soon after legislators arrive in Washington, many of them almost invariably find they can reap considerable rewards from oversight activity. 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 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.¹⁷² In the mid-1980s,

¹⁷¹ The agency (quite plausibly) assumes that if a recommendation is not implemented within four years of being made, it probably will not be. Interview with GAO Official # 1 (Washington, DC; January 2005).

¹⁷² See JAMES Q. WILSON, *BUREAUCRACY* (1990).

political scientists Mat McCubbins and Tom Schwartz introduced what has become an incredibly durable framework for thinking about legislative oversight of the bureaucracy.¹⁷³ 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 express profound concern with some aspect of bureaucratic activity.

Useful as it is from a political perspective, fire alarm oversight is precisely 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.¹⁷⁴

No doubt that congressional investigations often uncover important trends or problems in bureaucratic activity, whether such investigations are triggered by fire alarms or they arise from more pervasive police patrol methods. While it is 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 (or through a stratified random sample) tend to provide a biased sample. The results skew the picture of bureaucratic activity that emerges, 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 audited. Cases that are not reviewed at all don’t become the subject of any legislative, political, or public pressure. (2) Even when legislators and their staff choose to focus on a particular agency function, their oversight does not necessarily imply review of specific decisions. As with the audit bureaucracies, oversight

¹⁷³ See Mathew McCubbins and Thomas Schwartz, *Congressional Oversight Overlooked: Police Patrols vs. Fire Alarms*, 28 AM. J. POL. SCI. 165, 166-68 (1984).

¹⁷⁴ See, e.g., Light, *supra* note 43, at 42

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.

WHY AUDITS FACE POLITICAL AND ORGANIZATIONAL CONSTRAINTS

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

A. *Political Actors Have Polarizing Incentives*

Two important sets of actors who have stakes in the work of the executive branch may find their goals cut against audits of executive discretion. Officials in the executive branch (and their allies in the legislature) 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.¹⁷⁵

Here is the problem. Other things being equal, executive branch officials would be loath to part with discretion. Discretion helps authorities carry out the functions that they are expected to, like keeping threatening people out of the country, prevailing in 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*. Executive 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.¹⁷⁶ And some officials may simply crave power for its own sake. In short, other things being equal, executive authorities and their allies should be expected to seek more discretion, and less review. Unless those authorities have reason

¹⁷⁵ 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).

¹⁷⁶ Unless, of course their support of executive branch power is overwhelmed by incentives to support (for example) legislative power.

to limit their flexibility in order to demonstrate their competence,¹⁷⁷ 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.

What about legislators concerned about limiting executive power? Think first of the legislators who tend to distrust what the executive branch is doing. As noted earlier, legislators tend to lack incentives to audit rather than to rely on police patrol methods. *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.¹⁷⁸ A highly polarized debate has some benefits. It may galvanize support among certain constituencies. And opposition legislators (along with their allies in external interest groups), enthralled by the prospect of an optimal gamble, 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, organizations such as 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 assessment of the probability that they will prevail in advocating for the across-the-board standard. No doubt some determined advocates of more stringent judicial review would ground their commitments on their perception that courts are more politically insulated from legislative or executive pressures than the audit bureaucracies. They may laud courts' role in articulating the underlying nature of constitutional commitments, or to directly impose reforms on public bureaucracies through structural injunctions.¹⁷⁹ But in practice, these principled rationales may exacerbate a perception among critics of executive authority that measures short of substantially more stringent judicial supervision would yield little or no benefit.

In short, polarization among political advocates seeking maximal advantage in their efforts to expand or limit judicial review 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 a compromise. Conceptually, audits embody just such a compromise. Their architecture necessarily resonates most with observers who simultaneously worry about the drawbacks and benefits of greater review of

¹⁷⁷ For a discussion of the possibility that executive authorities could limit their discretion as a means of sending a costly signal of their to the public, *see* Cuéllar, *supra* note 67.

¹⁷⁸ 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.

¹⁷⁹ Because these observations raise valid concerns, the approach described here is designed to work in tandem with, rather than to entirely supplement, judicial review. *See supra* notes__.

executive discretion, and least with those for whom such a syncretic exercise seems unnecessary.

B. Institutional Inertia Locks In Existing Conceptions of Adjudication

To the extent that lawyers concerned with the exercise of discretionary powers view themselves as zealous advocates on behalf of individual clients, they may find little solace in a system that randomly selects cases for review. Even observers without individual clients to represent may naturally seek to focus attention on strategies to obtain direct relief for aggrieved individuals. This conception is likely to exist in some tension with the notion that more can be learned through reviewing fewer cases more thoroughly, even if the role of audits is to supplement rather than replace such review. Observers emphasize the value of adjudication as a recourse that should be available to, and provide a remedy, to similarly-situated parties. Judgments that don't provide a remedy may 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.¹⁸⁰

Audits of executive discretion do not conform to these assumptions. In a narrow sense, they randomly privilege some people – whose cases are selected for audits – and not others. They do not provide an obvious remedy, though it is certainly possible to forge a system that makes remedial contributions by affecting the ordinary course of judicial review.¹⁸¹ They seem, as a result, to be ill-fitting proxies for a persistent set of concerns that underlie the normative case for less deferential adjudication. It is undoubtedly true that constitutional provisions and values may require adjudication, and that many deficiencies in adjudication are best remedied through changes in adjudication.

Nonetheless, it seems equally clear that the prevailing conception of adjudication could unduly dampen interest in audits. It promotes the misleading 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 should attend to the internal and external procedures shaping the extent to which a specific agency decision becomes arbitrary.¹⁸²

Indeed, while *Vermont Yankee* may arguably limit courts' abilities to directly impose audit requirements through expansive interpretations of the Administrative Procedure Act, audits seem directly germane to the familiar procedural due process

¹⁸⁰ See Daryl J. Levinson, *Rights Essentialism and Remedial Equilibration*, 99 COLUM. L. REV. 857 (1999).

¹⁸¹ See *supra* Part I.c.

¹⁸² Cf. Lisa Schultz Bressman, *Beyond Accountability: Arbitrariness and Legitimacy in the Administrative State*, 78 N.Y.U. L. REV. 461 (2003).

balancing framework ordinarily traced to *Mathews v. Eldridge*.¹⁸³ By privileging the importance of accuracy-enhancing mechanisms, *Mathews* implies that (where protected interests are at stake) auditors' work could illuminate questions central to procedural due process analysis, such as the probability that individuals or groups will suffer erroneous deprivations under existing (as opposed to plaintiff-requested) procedures.¹⁸⁴ As they respond to conceptual and political objections, auditors (and audit supporters) may thus find solid ground from which to emphasize the potential intersections between the work of courts and auditors.

C. *Narrow Conceptions of Auditor Mission Prevail*

Like all bureaucracies, the GAO and IG Offices are also affected by prevailing conceptions of their mission found among its internal staff and leadership as well as external constituencies. 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 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.

Since the GAO and IG offices were created to serve as *auditors*, at least one factor shaping the priorities of these bureaucracies is rooted in organizational conceptions of how the role of an auditor should be defined. The legislators who created these bureaucracies and their successors may have long thought of these bureaucracies as a means of 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 hundred and thirty-three units, with annual funding in excess of \$20 billion, told GAO they had not received a financial audit during fiscal years 1974 through 1976.”¹⁸⁵

Politics has cemented the early focus on financial auditing and procurement.¹⁸⁶ The audit bureaucracies' legislative overseers expect them to demonstrate results, both in the sense of providing useful vehicles for legislators to achieve their own strategic objectives and in the more prosaic (and often overlapping) sense of detecting financial

¹⁸³ 424 U.S. at 319, 323-28.

¹⁸⁴ Indeed, the frequent relative absence of an empirical basis informing courts' determinations of what process is due in *Mathews*-like cases raises the question of how courts are currently making these determinations, and how these might be best evaluated.

¹⁸⁵ See Light, *supra* note ___, at 42.

¹⁸⁶ *Id* at 43-45.

improprieties that, where rectified, produce additional government revenues.¹⁸⁷ In some ways, these constraints are a reflection of legislators' own aforementioned incentives. Legislators can nudge or even force these agencies to undertake work the lawmakers desire. Given these realities, it is not surprising that the audit bureaucracies so consistently seem to embrace the fire alarm (and, to a more limited extent, the police patrol) approach associated with legislative oversight. 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.¹⁸⁸

On occasion, legislators may find audit bureaucracies useful to generate publicity and promote policy objectives not related to the advancing the ubiquitous mantra of eliminating waste, fraud, and abuse.¹⁸⁹ Nonetheless, the legislative requests that drive a considerable proportion of the GAO's work (and probably some of the work of Inspector General's offices) appear to reflect considerable resilience in the extent to which the audit bureaucracies are viewed as experts in investigating financial management: who spent what funds, why government vehicles were used for that trip, or why these employees were asked to work on some questionable task. The same may be largely true of managers and officials within the agencies themselves. Many Inspectors General have a background in financial management or accounting, as do a considerable proportion of staff at the GAO.

The audit bureaucracies are not entirely devoid of flexibility. It would almost certainly prove misguided to see the political and organizational constraints on audit bureaucracies as insurmountable, or their mission as entirely static. Even if they were, there is likely more than just legislative pressure at work in determining the audit bureaucracies' agenda. Recurring disagreements among legislators almost certainly leave the audit bureaucracies with a measure of discretion. Existing law already provides both the GAO and IG offices with broad jurisdiction to audit executive discretion. Over the last few decades, the audit bureaucracies have used that jurisdiction to generate (often at congressional request) the broad analyses of management practices and administrative priorities in public bureaucracies that, together with the aforementioned financial audits, constitute the bulk of their work.

Still, change is unlikely to come easily. Whether the audit bureaucracies assume the responsibility for auditing executive discretion or a separate auditing authority is created, greater use of that jurisdiction to effectively audit executive discretion depends in large measure on demand from politically significant constituencies such as legislators

¹⁸⁷ See, e.g., Testimony of the Hon. Pete Sessions, *Seven Years of GPRA: Has the Results Act Provided Results?* (July 20, 2000), avail. at 2000 WL 1211285 (F.D.C.H.) (“GAO has identified waste, fraud and abuse in government programs in amount of over \$800 Billion in the last 5 years.”).

¹⁸⁸ See, e.g., Anne Joseph, *The Politics of Auditing*, Unpublished Ph.D. Dissertation, Political Economy and Government Program, Harvard University (2001).

¹⁸⁹ See Light, *supra* note 43, 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.”).

and organized interest groups. Although change in their agendas is not impossible, the preceding discussion shows it to be unlikely. If political forces do not directly foster change then the fate of audits depends on the prospects for a degree of autonomy among the audit bureaucracies themselves.¹⁹⁰ Despite appearances to the contrary, it is neither necessary nor sufficient merely to enhance the formal independence of the audit bureaucracies. As with minority-protecting institutional rules like the Senate filibuster, legal changes in an agency's formal autonomy can invariably be undone unless the agency has managed to enshrine the notion (among politically relevant elites or the larger public) that its autonomy should be protected.¹⁹¹ The development of public bureaucracies during the last century provides ample evidence of agencies responding reflexively to political demands, but also of agencies forging coalitions and manipulating their political environment to become founts of major policy innovation as the nation evolved.¹⁹² At the Postal Service, the Forest Service, and the Food and Drug Administration, agency leaders empowered by having successfully led transformations of staff selection and promotion paths became determined policy innovators by cultivating support among professional elites and the mass public.¹⁹³ As the audit bureaucracies themselves evolve, their leaders may gradually awaken to the simple realization that they are capable of acquiring some measure of autonomy along these lines, overcoming political constraints by cultivating reputations for relative impartiality and technical competence. In the process, the audit bureaucracies remain capable of playing a unique and exceedingly valuable part in public life by auditing executive discretion. Their greatest legacy may lie in steadily forging coalitions that would make it possible for them to play that role.

CONCLUSION

The inexorable logic of executive discretion destines it to carry risks as well as rewards. Pervasive discretion lets government protect the environment, prosecute serial

¹⁹⁰ Cf. DANIEL CARPENTER, *THE FORGING OF BUREAUCRATIC AUTONOMY* (2002). The auditor's capacity to undertake audits in a manner that embodies some principled standard therefore depends on two critical factors. The first is whether the auditor manages to create a sufficient degree of autonomy from the ordinary political pressures that might keep the entity from being purely a creature of the legislature (or the executive). As Carpenter notes, this tends to turn on the capacity of the agency to forge ties to policymaking elites that support its mission and favorable impressions among the mass public (which suggests that, in a sense, it is these segments of the public guarding the proverbial guardians). The second is whether – given such autonomy – the agency's leaders and staff sufficiently value their information-generating mission to expend the necessary effort to generate material information.

¹⁹¹ An example from New Zealand is instructive. An ordinary parliamentary statute set up independent redistricting commissions. By its terms, the statute can only be modified by a 75% vote of Parliament. But because the redistricting statute itself could be repealed or supplanted by a majority vote in Parliament, the supermajority provision is essentially only advisory. Political observers indicate it would nonetheless be immensely challenging politically to repeal or change the statute. See Christopher Elmendorf, *Representation Reinforcement Through Advisory Commissions: The Case of Election Law*, 80 N.Y.U. L. Rev. 1366, 1386-90 (2005). In a similar sense, the formal legal autonomy of an agency such as the Federal Reserve Board is best understood (in the absence of compelling theory and data to the contrary) as a reflection rather than a cause of political circumstances (sometimes critically shaped by the agency itself) promoting an agency's autonomy.

¹⁹² See Carpenter, *supra* note __.

¹⁹³ *Id.* See also Daniel Carpenter, *The Political Crystallization of an Organizational Identity: Pharmaceutical Regulation at the FDA, 1947-1963*, Working Paper: Department of Government, Harvard University (2005)(on file with author).

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 is known about organizations, the people who run them, and the complicated legal mandates entrusted to them.

This article considered the implications of these facts in light of two other realities. Judicial review does not constrain the exercise of many forms of executive discretion that would nonetheless almost certainly benefit from some external review. Regulatory and prosecutorial enforcement decisions are among the most cogent examples. And we do next to nothing to audit how that discretion is used, despite the presence of compelling reasons to think that executive branch officials will have a relentless tendency to frequently misuse that discretion. Because some discretionary actions can signal competence and resolve to naïve observers among the mass public, executive officials may have an incentive to use their discretion to create favorable impressions. Some officials or their employees may be far less subtle and engage in willful misconduct that is unlikely to be detected. Even the most noble officials and organizations may have a harder time learning without external mechanisms to systematically review and critique their work.

When policymaking elites and organized interests discuss the costs and benefits of executive discretion, they tend to respond by fueling a familiar debate about the value of greater judicial scrutiny of executive discretion. While this article does not dismiss the value of such greater scrutiny, particularly in the provision of discrete remedies for aggrieved individuals and groups, 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 recognize that a substantial dimension of the problems associated with policing executive discretion involve *information*. Information is what impels the case for audits, which in turn hold the promise of severing 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. Indeed, 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.

On the other hand, 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 will tend to be 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. In the course of navigating the imperfect world in which law actually operates, the possibility of auditing executive discretion can be treated as a problem of institutional design. The discussion thus engendered can encompass questions about who should audit, how large samples should be, what standard should be used, or what should be the universe of cases to audit.

But a rich discussion of the means and methods for auditing executive discretion may serve a more immediate function. As a thought experiment, audits help elucidate the conceptual murkiness of many arguments for executive discretion and emphasize the underappreciated importance of audit bureaucracies such as the GAO and the Inspectors General. 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. If executive discretion is to be defended coherently against audits, the case for it must necessarily transcend the reflexive critiques of judicial supervision or the veneration of an amorphous, ill-defined political process. Instead, the case for insulating discretionary decisions from audits must provide a uniquely eloquent exaltation of the value of public ignorance.