



11-1-2006

Auditing Executive Discretion

Mariano-Florentino Cuellar

Follow this and additional works at: <http://scholarship.law.nd.edu/ndlr>

Recommended Citation

Mariano-Florentino Cuellar, *Auditing Executive Discretion*, 82 Notre Dame L. Rev. 227 (2006).

Available at: <http://scholarship.law.nd.edu/ndlr/vol82/iss1/6>

This Article is brought to you for free and open access by NDLScholarship. It has been accepted for inclusion in Notre Dame Law Review by an authorized administrator of NDLScholarship. For more information, please contact lawdr@nd.edu.

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, the 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 associated with limits on judicial review. Yet judicial review's evisceration as a tool to restrain certain forms of discretion also makes it easier for some officials to promote appealing political impressions by subtly manipulating decisions, for others to engage in outright malfeasance, and for still others to simply fail to correct mistakes. Reliance on judicial review to generate information about executive discretion makes it difficult to address these concerns because courts routinely define much of their work in terms of applying the same standard of deference to every case in a particular class, limiting possibilities to increase the stringency of review in some policy domains without making the costs allegedly prohibitive. As a conceptual alternative for monitoring executive discretion, this Article develops a framework akin to that employed by courts engaged in the "sample adjudication" of class action and government fraud cases. It relies on the possibility of systematically auditing samples of discretionary decisions and making those

© 2006 Mariano-Florentino Cuéllar. Individuals and nonprofit institutions may reproduce and distribute copies of this Article in any format, at or below cost, for educational purposes, so long as each copy identifies the author, provides a citation to the *Notre Dame Law Review*, and includes this provision and copyright notice.

* Associate Professor and Deane F. Johnson Faculty Scholar, Stanford Law School; Faculty Affiliate, Stanford Center for International Security and Cooperation. I appreciate helpful comments on earlier versions of this project from Glenn A. Fine, Daniel Sutherland, Mark Kelman, Pamela Karlan, Barbara Fried, Robert Weisberg, Deborah Hensler, Rachel Barkow, Mark Lemley, Jayashri Srikantiah, Laura Donohue, Mitchell Polinsky, Lawrence Friedman, Michael Barr, Sallyanne Payton, Sally Katzen, Jeff Strnad, Chris Elmendorf, Bruce Owen, and Joseph Grundfest. The project also benefited from the feedback of 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. This Article is dedicated to Pistachio.

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 value, such audits are nonetheless almost never undertaken by existing federal audit bureaucracies, nor does the legislature seem to conduct them in connection with oversight hearings. This Article discusses the dynamics working against these audits, explains how auditing may nonetheless occasionally prove to be politically viable, and concludes by emphasizing the importance of greater sensitivity to institutional complexities in recurring debates about the merits of executive discretion.

INTRODUCTION	229
I. THE CASE FOR AUDITING EXECUTIVE DISCRETION	236
A. <i>Definitions</i>	236
B. <i>Examples of Routine Executive Discretion</i>	240
1. Example # 1: Generally Unreviewable Executive Decisions	241
2. Example # 2: Highly Deferential Review	243
a. Asset Freezing	244
b. Environmental and Occupational Safety Administrative Compliance Orders	247
C. <i>How Discretion is Managed Through Variations in the Stringency of Traditional Judicial Review</i>	249
D. <i>Why Audits Could Serve as a Valuable Supplement to Traditional Judicial Review</i>	252
1. The Role of Audits in Overseeing Complex Organizations	252
2. The Risks of Tolerating the Amount of Discretion Associated with Traditional Judicial Review	258
a. Arguments for Expansive Discretion and Their Contingent Nature	259
b. The Risks Associated with Expansive Discretion	261
3. A Social Welfare-Oriented Case for Auditing Executive Discretion	265
4. How Audits Would Fare in a Politically Fragmented World	274
II. WHY THE INSTITUTIONAL DESIGN PROBLEMS INHERENT IN AUDITS ARE MANAGEABLE	278
A. <i>What Executive Decisions Should be Audited?</i>	279
B. <i>What Sample Should Be Used?</i>	282

C. <i>Which Organization Should Audit and What Standard Should It Apply?</i>	285
D. <i>What Should Depend on the Outcome of an Audit?</i>	288
III. WHY REFORMS CONTRAST WITH THE STATUS QUO	290
A. <i>Audit Bureaucracies Do Not Ordinarily Perform Audits</i>	292
B. <i>Neither Do Legislators</i>	296
IV. WHY AUDITS FACE CONSTRAINTS	299
A. <i>Political Actors Have Incentives to Undermine Audits</i>	299
B. <i>Institutional Inertia May Exaggerate the Value of Adjudication</i>	302
C. <i>Audit Bureaucracies Define Their Mandates Narrowly</i>	304
CONCLUSION	308

INTRODUCTION

The problem is a familiar one. Most laws are enforced to some degree by executive bureaucracies. No legal system can ever vanquish their discretion. Even if doing so were possible, it would be madness to squeeze out every last drop of human judgment from the law's application. Then again, too much discretion breeds its own madness. Sensitive to this predicament, legislators and judges subject a range of government activities to elaborate legal constraints.¹ Procedural formalities and external court review—ostensibly designed to balance the risks and rewards of lodging public power in bureaucratic organizations—therefore epitomize familiar legal tasks such as criminal trials and regulatory rulemaking proceedings.

But not every task. Despite the presence of these constraints in some domains,² the hallmark of many executive decisions often

1 For insightful discussions of how concerns about discretion led to the enactment of much of modern administrative law, see McNollgast, *The Political Origins of the Administrative Procedure Act*, 15 J.L. ECON. & ORG. 180, 195–201 (1999); Peter L. Strauss, *Changing Times: The APA at Fifty*, 63 U. CHI. L. REV. 1389, 1399–1400 (1996).

2 See, e.g., *Citizens to Pres. Overton Park, Inc. v. Volpe*, 401 U.S. 402, 411 (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. *Id.* at 410. But review remains either unavailable or fairly cursory for a massive range of discretionary decisions involving national security, foreign policy, immigration, domestic regulatory enforcement, public benefits, and investigation or prosecution. See KENNETH CULP DAVIS, *DISCRETIONARY JUSTICE* 151–55 (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, 317–19 (2002); David Cole,

proves to be nearly unfettered discretion.³ 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. These reservoirs of discretion persist even in settings where judicial review of executive branch action is considered a central tool, one that not only resolves individual claims but prevents systemic mistakes or abuses.⁴ Yet the presumed benefits of discretion have led to a vigorous doctrinal and policy debate about the proper scope and stringency of such review,⁵

Judging the Next Emergency: Judicial Review and Individual Rights in Times of Crisis, 101 MICH. L. REV. 2565, 2576–77 (2003); Lisa Schultz Bressman, *Beyond Accountability: Arbitrariness and Legitimacy in the Administrative State*, 78 N.Y.U. L. REV. 461, 473–74 (2003); James E. Pfander, *Article I Tribunals, Article III Courts, and the Judicial Power of the United States*, 118 HARV. L. REV. 643, 653 (2004).

3 See *infra* Part I.B.

4 See, e.g., *City of Chi. v. Int’l Coll. of Surgeons*, 522 U.S. 156, 184 (1997) (summarizing previous interpretation of federal jurisdictional statute by emphasizing the Court’s conclusion that “Congress meant to hold federal agencies accountable by making their actions subject to judicial review”); Rebecca L. Brown, *Accountability, Liberty, and the Constitution*, 98 COLUM. L. REV. 531, 535 (1998) (emphasizing the alleged role of judicial review in promoting accountability); 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.”). 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, for example, Michael Asimow, *The Scope of Judicial Review of Decisions of California Administrative Agencies*, 42 UCLA L. REV. 1157, 1239 (1995); Jeffrey J. Rachlinski & 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, 509–12 (2002); 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, 333 (1976).

5 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 2, 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

and in some cases—such as those where a choice is committed by law to agency discretion under federal administrative law—to almost no review at all.

This Article challenges the terms of that debate. In the process, it raises questions about commonly asserted justifications for insulating executive discretion from external review. The argument demonstrates how the traditional paradigm of judicial review, despite its enduring value, sometimes ill serves the goals of helping bureaucratic organizations learn from their failures and avoid political pressures endangering their missions.⁶ Because complex public bureaucracies throughout government are increasingly—and perhaps inevitably—the custodians of discretionary legal authority that can be abused,⁷ the problems arise both in national security and domestic regulatory contexts—domains that have been traditionally treated separately but increasingly blur.⁸ The Article then explains how a government agency

to be necessary should be properly confined, structured, and checked.”); Cole, *supra* note 2, 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, 1420 (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 & Adrian Vermeule, *Accommodating Emergencies*, 56 STAN. L. REV. 605, 644 (2003) (“Judicial scrutiny can only interfere with forceful executive action.”); Ruth Wedgwood, *Al Qaeda, Terrorism, and Military Commissions*, 96 AM. J. INT’L L. 328, 332–35 (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 *infra* Part III.

6 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 *infra* Part II.C (discussing the relevance of the argument to court review of individual cases); Part II.D (discussing the implications of the argument for how political actors deliver remedies to aggrieved individuals or groups).

7 The focus here is primarily on the type of routine executive 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, 837–38 (1985) (holding that in the absence of a 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). For a discussion of routine executive discretion, see *infra* notes 29–34 and accompanying text.

8 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.”).

can perform quasi-judicial *audits* of discretionary decisions, akin to the sample audits occasionally employed by courts in class actions and government fraud cases.⁹ If an auditing agency overcomes the relevant political barriers and conceptual challenges,¹⁰ it can fill crucial gaps left by existing mechanisms to generate information about executive branch performance.

Part I begins by describing how judicial review plays a prominent role in generating information about executive branch decisions. The picture of executive discretion that emerges in such a system 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 that the issue is the fate of individuals designated 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.¹¹ Greater stringency of review presumably reduces the probability that someone would be improperly labeled an enemy combatant. At the

9 See *infra* text accompanying notes 105–07.

10 For a discussion of the proverbial “guarding the guardians” problem, see *infra* notes 211–13 and accompanying text.

11 Compare *Webster v. Doe*, 486 U.S. 592, 602–04 (1988) (finding CIA director's power to fire employees on national security grounds committed by law to agency discretion), with *Klamath-Siskiyou Wildlands Ctr. v. Bureau of Land Mgmt.*, 387 F.3d 989, 997 (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 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 cited represent extraordinarily different contexts, and the kinds of discretion involved in the decisions are also different. The point is that 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 courts' 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.

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 decisions.¹² 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 and legislatures tend to vary the stringency of review governing a given pool of potential cases.

But the assumption that courts should apply the same degree of review stringency to every present and future case in a particular class 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.¹³ Increasing the stringency of review for a single decision appears difficult, if not impossible, without sharply increasing costly burdens on courts and the government. Even if one assumes that existing rules governing stringency of review reflect 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 routinely 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.¹⁴ Government agents select a subset of plants to inspect or accounting records to scrutinize.¹⁵ This 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 relative absence from discussions of how to constrain government discretion, audits of this

12 See, e.g., Brief for the Respondents at 11–12, *Hamdi v. Rumsfeld*, 542 U.S. 507 (2004) (No. 03-6696), 2004 WL 724020 (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"); Reply Brief for the Petitioner at 4, *Heckler v. Chaney*, 470 U.S. 821 (1985) (No. 83-1878), 1984 WL 566059 ("[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)).

13 Moralistic intuitions about the importance of 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.

14 See *infra* text accompanying notes 109–10.

15 See *infra* Part I.D.1.

kind are familiar from 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 what judicial review could already achieve, either at existing or higher levels of stringency. Under quite reasonable assumptions, it is preferable from a social welfare perspective to use new resources at the margin for auditing instead of using them to expand deferential judicial review. Audits may even be valuable in a world where politicians are divided about what constitutes effective performance—though implementing audits under such conditions will tend to prove difficult.

Admittedly, the case for audits is a qualified one. Political constraints may occasionally check discretionary abuses. As will become clear, not all types of executive discretion should be placed in the same analytical category. The case for auditing may be stronger in the case of discrete applications of legal authority to individuals or groups than in the case of broad policy decisions that courts have historically excluded from review under the political question doctrine.¹⁶ Nonetheless, the absence of audits for routine discretionary decisions almost certainly diminishes the capacity of legislators and the public to detect executive branch manipulation,¹⁷ dampens 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

16 See *infra* Part I.A.

17 Of course, the mere creation of some auditing system does not automatically solve organizational learning and accountability problems. As Part II explains, 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 Article. 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 generally MICHAEL POWER, *THE AUDIT SOCIETY: RITUALS OF VERIFICATION* 123 (1999) (“The audit society is a society that endangers itself because it invests too heavily in shallow rituals of verification at the expense of other forms of organizational intelligence.”).

in statutory or executive mandates) what standards are actually supposed to govern executive discretion. 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.¹⁸ 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 broadens the focus beyond judicial review. Obviously, formal judicial review is not the only means through which executive power may be constrained. Other factors include some combination of the press, audit bureaucracies such as the Government Accountability Office (GAO) and Inspectors General (IG), legislators, and even organized interests and social movements drawing support from among the mass public. As the analysis below explains, formal review mechanisms—such as audits or judicial review—are often critical in galvanizing subsequent interest from legislators, the media, organized interests, and the public.

Moreover, it turns out that 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.¹⁹ As Part IV indicates, the relative absence of sampling techniques in the review of executive branch legal decisions may arise from a conceptual blurring of the direct-remedy and information-producing functions of review, and probably leads 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 bureau-

18 See *infra* Part I.D.1 (discussing the potential for dynamic interaction between audits and judicial review).

19 As Part I explains, this means that the audit bureaucracies generally appear not to: (1) take random samples of decisions in a legal domain involving decisions directly affecting individuals or groups, (2) assess decisions in accordance with a defensible pre-defined standard, and (3) release the results publicly.

cracies, drawing their attention to procurement fraud and similarly tangible examples of waste. Furthermore, 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. The status quo also makes it easier for advocates of executive power to confound different kinds of arguments for limiting external review—those based on concerns about the practical burdens associated with judicial review, and those based on structural, separation of powers concerns associated with doctrines such those governing the review of so-called political questions. In response, this case study and thought experiment aims to clarify matters by focusing on three recurring challenges that arise in connection with existing (sub-optimal) institutional responses: (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. To address those challenges, this Article seeks to broaden the scope of potential solutions, and to shed light on the forces that shape public perceptions of whether those solutions even exist.

I. THE CASE FOR AUDITING EXECUTIVE DISCRETION

A. *Definitions*

Discretion can be defined as the extent of legal flexibility to use government power vested in executive branch officials—including, but not limited to, power over personnel, budgets, information, and legally-sanctioned coercive authority to affect the world.²⁰ Govern-

²⁰ Among the nuances distinguishing different types of discretion, one might recognize a distinction between the actual, *de jure* amount of flexibility the law permits (as in the examples of routine executive discretion), and the *de facto* amount of discretion that results from the fact that review of a given decision is limited or almost

ment 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. Consequently, distinctions in the amount of executive discretion are relative. 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. Outside the context of the political question doctrine, lawyers, judges, and scholars focusing on federal regulatory activity have understood that baseline to include court review of executive action.²¹

At the same time, 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, Inc. v. Volpe*²² firmly establishes close judicial scrutiny of a discretionary decision as a presumptive means of policing executive decisions,²³ other cases in 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 reasons of constitutional structure, entirely to the elected branches.²⁴ In contrast to this “political question discretion,” a some-

entirely absent (as may occur in the context of police searches involving situations that do not give rise to criminal proceedings and therefore are not policed through the exclusionary rule). Although the primary focus of the argument is on the former type of discretion, the present analysis also has implications for the latter—because expanding the range of available review mechanisms may change policymakers’ and legal interpreters’ willingness to impose review that could change the extent of de facto discretion.

21 See, e.g., *Abbott Labs. v. Gardner*, 387 U.S. 136, 141 (1967) (relying on presumption of reviewability in federal administrative law). For examples from the states, see *State Dep’t of Fish & Game v. Meyer*, 906 P.2d 1365, 1370 (Alaska 1995) (acknowledging that final agency action is presumed reviewable); *Minn. Pub. Interest Research Group v. Minn. Envtl. Quality Council*, 237 N.W.2d 375, 379 (Minn. 1975) (recognizing presumption of reviewability for state administrative decisions); *Pisano v. Shillinger*, 835 P.2d 1136, 1138–39 (Wyo. 1992) (presuming reviewability of state administrative action). For a discussion of how the institutional design of judicial review in the European Union provides for the review of administrative action, see generally Jurgen Schwarze, *Judicial Review in EC Law—Some Reflections on the Origins and the Actual Legal Situation*, 51 INT’L & COMP. L.Q. 17, 22–27 (2002).

22 401 U.S. 402 (1971).

23 *Id.* at 410.

24 See, e.g., *Marbury v. Madison*, 5 U.S. (1 Cranch) 137, 170 (1803). As I acknowledge below, there may indeed be 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

what different form of discretion exists that could be described as "routine executive discretion." It is this form of discretion that was the subject of another canonical case, *Heckler v. Chaney*,²⁵ where the Court considered the appropriate degree of judicial scrutiny of a regulatory agency's discretionary enforcement decisions²⁶ and emphasized the value of limiting judicial intervention in such matters by analogizing regulatory enforcement discretion to prosecutorial discretion in the criminal context.²⁷ 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,²⁸ preserving the type of bureaucratic discretion at issue in *Chaney* and similar cases from judicial intervention raises the risk that executive authorities will behave in an arbitrary fashion.²⁹

It will come as no surprise that the arguments for discretion in *Chaney* and similar cases are routinely bolstered by assertions about the value of bureaucratic expertise. In *Panama Canal Co. v. Grace Line, Inc.*,³⁰ for example, the Court considered the claims of American shipping companies demanding that new tolls be set on the Panama Canal.³¹ The key passage of the Court's reasoning applied the following now-familiar logic:

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.

25 470 U.S. 821 (1985).

26 *Id.* at 831.

27 *Id.* at 834-35. The analogy bolstered the Court's case for limiting judicial interference in the *Chaney* 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, 364 (1978) ("In our system, so long as the prosecutor has probable cause to believe that the accused committed an offense defined by statute, the decision whether or not to prosecute, and what charge to file or bring before a grand jury, generally rests entirely in his discretion.").

28 *See, e.g., Bressman, supra* note 2, at 473-74 (noting that judicial review is effective as the main protection against abuse of executive discretion).

29 *See infra* Part I.D.2 for a discussion of those risks. Although cases like *Chaney* may seem less troubling to some because they involve a government decision to *restrain* from acting, such a decision still counts as action under the Administrative Procedure Act, *see* 5 U.S.C. § 551(13) (2000), and is potentially no less coercive to individuals whose desired outcomes depend on government action. Part II, *infra*, discusses how the practical problems of reviewing executive decisions *not* to act may be solved through alternatives to traditional judicial review.

30 356 U.S. 309 (1958).

31 *Id.* at 310.

[T]he present conflict rages over questions that at heart involve problems of statutory construction and cost accounting: whether an operating deficit in the auxiliary or supporting activities is a legitimate cost in maintaining and operating the Canal for purpose of the toll formula. These are matters on which experts may disagree; they involve nice issues of judgment and choice, which require the exercise of informed discretion.³²

Even when lawyers advocating on behalf of executive power echo the Court in extolling the value of such *Chaney*-type executive discretion, they nonetheless tend to implicitly accept a baseline state of the world where courts play a significant role in reviewing government action.³³ 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).³⁴ 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 diverge on how easy it should be for a court to impose a labor injunction or to convict a defendant. But if they fail in persuading the legislature to water down the substantive standard that applies, advocates of discre-

32 *Id.* at 317 (citing *New York v. United States*, 331 U.S. 284, 335 (1947)).

33 *See, e.g.*, Brief for the Respondents, *supra* note 12, at 25. 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. (citations omitted) (quoting *Dep't of the Navy v. Egan*, 484 U.S. 518, 530 (1988)).

34 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 is 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) (recognizing differences in the “form, function, and degree of reciprocal engagement” among legal systems).

tion 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 fine will be subsequently invalidated.³⁵ These realities imply that we can measure the *benefits* of executive discretion against a baseline of relatively intrusive judicial review.³⁶ We can also assess the potential benefits of alternative means of reviewing routine discretionary actions—such as random audits of those decisions—by comparing those alternatives to the existing framework of traditional judicial review.

B. *Examples of Routine Executive Discretion*

As a prelude to understanding the case for an alternative means of review, consider two categories of routine executive discretion involving bureaucratic flexibility to use legal authority affecting individuals and groups. Though observers may argue about the prescriptive merits of lodging so much discretion in the bureaucracy, neither category reflects a judicial decision to exercise restraint on account of the political question doctrine or structural separation of powers ratio-

35 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* 147–71 (1990) (suggesting that the National Highway Traffic Safety Administration’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); Brandice Canes-Wrone, *Bureaucratic Decisions and the Composition of the Lower Courts*, 47 AM. J. POL. SCI. 205, 212–13 (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 fourteen percent); Thomas O. McGarity, *The Courts and the Ossification of Rulemaking: A Response to Professor Seidenfeld*, 75 TEX. L. REV. 525, 528 (1997) (discussing the impact of the “ossification” of rulemaking, where judicial review, among other factors, shapes agencies’ willingness to use regulatory authority).

36 Moreover, there is no good reason to expect that courts or legislators routinely strike the proper balance (under almost any defensible definition of “proper”) when they police discretion outside the context of separation of powers and the political question doctrine. My claim is this: at a minimum, *Chaney*-type discretion directly affecting individuals and justified on consequentialist grounds concentrates great power in executive authorities and raises questions about the appropriate extent of bureaucratic flexibility that should be permitted.

nales.³⁷ Instead, what these examples tend to reflect is a pragmatic calculus in judicial opinions, or encompassed by legislative enactments subject to judicial interpretation.

1. Example # 1: Generally Unreviewable Executive Decisions

A small but important class of executive decisions are subject only to the barest degree of judicial review. In federal administrative law, certain activities are described (in a somewhat tautological fashion) as being committed by law to agency discretion. *Chaney* itself is a classic example. Death row inmates were thwarted in their quest to force the FDA into blocking state prison authorities from using drugs for an unapproved purpose consisting of execution.³⁸ But analysts have long understood how the case's larger implications lie beyond its unusual facts.³⁹ The case cements a barrier to review of bureaucratic decisions involving the nonenforcement (or enforcement) of regulatory mandates. As the situation for the death row inmates readily illustrates, decisions involving how to enforce legal mandates (e.g., involving prohibitions on the unapproved use of pharmaceutical products) can have effects on some individuals and groups as coercive as decisions regarding the substance of a regulatory or statutory mandate itself (e.g., for what uses a product is approved).

Yet the case also hints at the practical problems that would ensue from a full judicialization of agency enforcement discretion. It borders on madness for courts (or legislatures) to allow protracted litigation whenever a party is aggrieved by its decision to enforce legal mandates in a manner other than how the litigant believes the agency should.⁴⁰ As the *Chaney* Court readily appreciated, a similar problem

37 The invocation of such doctrines could also be scrutinized and subjected to criticism, but such an inquiry is beyond the scope of the present Article.

38 *Heckler v. Chaney*, 470 U.S. 821, 823–28 (1985).

39 See, e.g., Ronald M. Levin, *Understanding Unreviewability in Administrative Law*, 74 MINN. L. REV. 689, 692 (1990); Cass R. Sunstein, *Reviewing Agency Inaction After Heckler v. Chaney*, 52 U. CHI. L. REV. 653, 654 (1985). For an insightful pre-*Chaney* perspective on reviewability, see generally Harvey Saferstein, *Nonreviewability: A Functional Analysis of "Committed to Agency Discretion,"* 82 HARV. L. REV. 367 (1968).

40 See *Chaney*, 470 U.S. at 831–32 (“[A]n agency decision not to enforce often involves a complicated balancing of a number of factors which are peculiarly within its expertise. . . . An agency generally cannot act against each technical violation of the statute it is charged with enforcing. The agency is far better equipped than the courts to deal with the many variables involved in the proper ordering of its priorities.”). From the Court’s perspective, matters would have been different if legislators explicitly crafted statutes to require such review. See *id.* at 832–33 (“[T]he decision is only presumptively unreviewable; the presumption may be rebutted where the substantive statute has provided guidelines for the agency to follow in exercising its enforcement

arises in the context of prosecutorial discretion of federal, state, and local criminal justice authorities.⁴¹ And a similar conclusion is reached by courts in that context: the system would break down if stringent court review of prosecutorial charging decisions (or declinations) became anything other than exceedingly rare—available only in exceptional circumstances, if at all.⁴² Indeed, courts have recurrently found that prosecutors harbor an inherent power to choose whom to charge with few (if any) judicially imposed constraints.⁴³

Courts are rarely entirely absent from reviewing discretionary actions. Judges and the legislators who shape the relevant statutory mandates may pull back from the brink of concluding that they will reject any possible claim associated with an ostensibly unreviewable decision. The law says, for example, that the CIA Director has authority to fire employees for being national security risks.⁴⁴ 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).⁴⁵ The decision of the CIA Director is, however, subject to limited constitutional review.⁴⁶ Even when all or most alternative review is precluded, courts can restrain egregious government conduct of some kinds, such as

powers.”). Absent some exceptional circumstances, they would be loath to do so. *Cf.* *Bowen v. Mich. Acad. of Family Physicians*, 476 U.S. 667, 675 (1986) (bolstering a conclusion that Congress had deliberately precluded judicial review of the amount of Medicare Part B claims in order to “‘avoid overloading the courts’ with ‘trivial matters,’ a consequence which would ‘unduly ta[x]’ the federal court system with ‘little real value’ to be received by participants in the program” (quoting *United States v. Erika, Inc.*, 456 U.S. 201, 210 n.13 (1982))).

41 *See Chaney*, 470 U.S. at 831 (“This Court has recognized on several occasions over many years that an agency’s decision not to prosecute or enforce . . . is a decision generally committed to an agency’s absolute discretion.”); *id.* at 834 (concluding that the court of appeals erred in finding inapplicable the “principle of absolute prosecutorial discretion”).

42 *See, e.g., United States v. Armstrong*, 517 U.S. 456, 464 (1996); *United States v. Nixon*, 418 U.S. 683, 693 (1974); *Vaca v. Sipes*, 386 U.S. 171, 182 (1967). For an insightful doctrinal review, see Lara Beth Sheer, *Prosecutorial Discretion*, 86 *Geo. L.J.* 1353 (1998).

43 *Armstrong*, 517 U.S. at 464. 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 2, at 209 n.21.

44 50 U.S.C.A. §§ 403–404a(e) (West Supp. 2006).

45 *Webster v. Doe*, 486 U.S. 592, 602–04 (1988).

46 A variety of practical problems for the claimant, beginning with discovery and exacerbated by courts’ tendency to defer to executive branch national security decisions, make it doubtful this avenue for review would exert much of an effect on the behavior of the agency’s leadership.

those that might give rise to viable constitutional tort claims.⁴⁷ But the sliver of review available in the context of federal constitutional claims (or similar claims based on state law) does not change an enduring reality: bureaucratic decisionmakers retain nearly unfettered freedom from review in a bewildering range of contexts directly affecting individuals and groups, involving regulatory enforcement, prosecutorial discretion, and national security authority. Defenders of existing limits may offer defensible rationales for not subjecting these domains to routine judicial review or its equivalent.⁴⁸ But the limits on review also engender a series of risks, discussed in more detail below.

2. Example # 2: Highly Deferential Review

Despite the fact that some decisions are formally committed to agency discretion, most of the time legislators consider regulatory actions too important to commit entirely to agency discretion. Instead they create procedural mandates—such as the Administrative Procedure Act (APA)⁴⁹—to restrain the scope of executive discretion. In still other cases, courts interpret constitutional provisions and statutory mandates to impose procedural obligations on executive bureaucracies.⁵⁰ While procedural requirements help determine the distribution of scarce political and legal resources among political and legal decisionmakers, at times the constraints they impose turn out to be milder than they first appear. The constraints then become part of a system of de facto discretion. Despite the formal availability of review, the meager extent to which decisions are actually scrutinized exacerbates certain risks—of mistakes, politically motivated self-dealing, and outright malfeasance—while often making it more difficult for external legislative or media-driven checks on executive discretion to operate.⁵¹ The following two examples—featuring the contexts of national security as well as domestic regulation—illustrate the dynamic.

47 See, e.g., Peter M. Shane, *Interbranch Accountability in State Government and the Constitutional Requirement of Judicial Independence*, LAW & CONTEMP. PROBS., Summer 1998, at 21, 25; sources cited *supra* note 2.

48 See Steven G. Calabresi & Saikrishna B. Prakash, *The President's Power to Execute the Laws*, 104 YALE L.J. 541, 570–99 (1994); Saikrishna B. Prakash & Michael D. Ramsey, *The Executive Power Over Foreign Affairs*, 111 YALE L.J. 231, 295–355 (2001); John Yoo, *War, Responsibility, and the Age of Terrorism*, 57 STAN. L. REV. 793, 810 (2004).

49 5 U.S.C. §§ 551–559, 701–706 (2000 & Supp. IV 2004).

50 See *Hamdi v. Rumsfeld*, 542 U.S. 507, 532 (2004) (plurality opinion).

51 Why this state of affairs may be problematic for dominant legislative coalitions without prompting them to fix it is discussed in Part IV.

a. Asset Freezing

The President and federal officials assisting him have powerful tools to regulate foreign economic activities. One such tool, involving designations under the International Emergency Economic Powers Act (IEEPA),⁵² 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.⁵³ 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.”⁵⁴ Second, the President must use this power only during an emergency.⁵⁵ 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.⁵⁶ It must pose a threat to the “national security, foreign policy, or economy of the United States.”⁵⁷ 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).⁵⁸ 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 (OFAC) blocks the assets of groups branded “Specially Desig-

52 50 U.S.C. §§ 1701–1706 (2000). See generally James J. Savage, *Executive Use of the International Emergency Economic Powers Act—Evolution Through the Terrorist and Taliban Sanctions*, CURRENTS: INT’L TRADE L.J., Winter 2001, at 28, 28–31 (describing the history and effect of IEEPA).

53 50 U.S.C. § 1701(a).

54 See *Global Relief Found., Inc. v. O’Neill*, 315 F.3d 748, 752–53 (7th Cir. 2002).

55 See 50 U.S.C. § 1701(a).

56 *Id.*

57 *Id.*

58 See generally Jason Luong, Note, *Forcing Constraint: The Case for Amending the International Emergency Economic Powers Act*, 78 TEX. L. REV. 1181, 1197 (2000) (“Congress has consistently acquiesced to the economic regulations enacted by the president under the IEEPA because of inadequate legislative drafting, lack of political will, and popular support for the most common of the IEEPA regulations—economic sanctions and embargoes.”).

nated Terrorist Organizations.”⁵⁹ Once a group becomes a specially designated terrorist organization, it loses control over its fate. The impact of the 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 \$50,000,⁶⁰ and willful violators are subject to criminal penalties, including up to twenty years imprisonment and fines of up to \$50,000 per violation.⁶¹ A similar designation also triggers severe criminal penalties punishing individuals for providing “material support” (including funds and in-kind economic contributions such as lodging) to designated terrorist organizations.⁶²

Take a closer look at how a court reviews the government’s designations. In the recent case, *Holy Land Foundation for Relief and Development v. Ashcroft*,⁶³ the State and Treasury departments used their delegated presidential IEEPA powers to freeze the assets of the Holy Land Foundation (HLF).⁶⁴ Court review of the blocking order considered whether it was “arbitrary and capricious” under the terms of the Administrative Procedure Act.⁶⁵ 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.⁶⁶ The district court, for example, emphasized the limited scope of its role.⁶⁷ 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”).⁶⁸ The district court conducted a careful examination of the record and uncovered “ample” evidence that

- (1) HLF has had financial connections 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

59 *Holy Land Found. for Relief & Dev. v. Ashcroft*, 333 F.3d 156, 160 (D.C. Cir. 2003).

60 50 U.S.C.A. § 1705(a) (West Supp. July 2006).

61 *Id.* § 1705.

62 18 U.S.C. § 2339 (2000 & Supp. III 2003).

63 333 F.3d 156.

64 *Id.* at 159–60.

65 *Id.* at 162.

66 *Id.*

67 *Id.*

68 *Id.* at 161.

Jerusalem office acted on behalf of Hamas; and (6) FBI informants reliably reported that HLF funds Hamas.⁶⁹

The D.C. Circuit upheld this determination on appeal.⁷⁰ 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.⁷¹ 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."⁷² That record may be a tremendously accurate compilation of the government's evidence. Or it may be patently misleading. Nothing requires the government to report evidence tending to cast doubt on its contentions.⁷³ Nor does the court interview the sources on which the record is based; thus court review is only as good as the record.

The flip side of this point is that court review will probably 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

69 *Id.* (quoting *Holy Land Found. for Relief & Dev. v. Ashcroft*, 219 F. Supp. 2d 57, 69 (D.D.C. 2002)).

70 *Id.* at 161-63.

71 *Id.* at 163 (applying these standards in reviewing the actions and record at hand).

72 *Id.* at 162.

73 Ironically enough, the *Holy Land* appeals panel suggested that the government's position was strengthened by the fact that "[t]here was no plausible evidence presented which showed that [ties to Hamas] had been severed." *Id.*

look totally arbitrary. Yet the record itself is based on decisions that are essentially immune from review.⁷⁴

b. 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 have some reason to think businesses are violating their general duty to provide a safe working environment, the agency can issue abatement orders and citations.⁷⁵ 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.⁷⁶ Under that statute, the Environmental Protection Agency (EPA) can issue an Administrative Compliance Order (ACO) on the basis of “any available information,” directing a regulated party—such as an electricity generation plant—

⁷⁴ Manipulation of the record, moreover, need not be conscious or explicit. A number of pressures could affect the 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).

⁷⁵ See 29 U.S.C. § 666(a)–(c) (2000).

⁷⁶ 42 U.S.C. §§ 7401–7671q (2000 & Supp. III 2003).

or state agency to comply with the Clean Air Act's requirements.⁷⁷ While the ACO does not allow the EPA to impose fines or other penalties directly, the order triggers provisions imposing civil or criminal penalties for violation of the order.⁷⁸ 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.⁷⁹ 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.⁸⁰ The Supreme Court declined to review this case, and instead—in a separate case—upheld a Ninth Circuit opinion holding ACOs to be final agency actions and reviewing them under the “arbitrary and capricious” standard.⁸¹

Both agencies therefore retain enormous power over when 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.”⁸² 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 bun-

77 *Id.* § 7413(a)(5).

78 *See id.* § 7413(b)–(c). 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, 197–99 (2005) (discussing the EPA's use of ACOs to enforce the Clean Air Act).

79 *See* 42 U.S.C. § 7413(b)(2).

80 *See* *Tenn. Valley Auth. v. Whitman*, 336 F.3d 1236, 1258 (11th Cir. 2003).

81 *See* *Alaska Dep't of Env'tl. Conservation v. EPA*, 540 U.S. 461, 496 (2004); *Alaska Dep't of Env'tl. Conservation v. EPA*, 298 F.3d 814, 822 (9th Cir. 2002), *aff'd* 540 U.S. 461.

82 42 U.S.C. § 7413(a)(5).

dle of substantive powers and penalties that could be acceptable to an enacting legislative and interest group coalition.

C. *How Discretion is Managed Through Variations in the Stringency of Traditional Judicial Review*

The preceding examples demonstrate how 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 operation of so-called "no-fly" lists impacting Americans' air travel activities,⁸³ the meager judicial scrutiny of discretion vested in government contractors engaged in quasi-official functions,⁸⁴ the enormous power federal officials wield (even in the wake of recent Supreme Court decisions on the subject) in designating enemy combatants,⁸⁵ and even in the myriad decisions governing federal procurement predicated on the exercise of government officials' legal discretion.⁸⁶ How, then, should courts and legislatures respond to perceived inadequacies in the degree of external scrutiny for these decisions?

What courts and legislatures have *not* done is to move towards maximally limiting discretion, by subjecting every decision to the type of stringent review associated with *de novo* fact-finding. This courts and legislatures are understandably loath to do in domains where they perceive relatively broad executive discretion to have benefits, and where they perceive judicial resources to be scarce. 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

83 See Justin Florence, Note, *Making the No Fly List Fly: A Due Process Model for Terrorist Watchlists*, 115 YALE L.J. 2148, 2157–59 (2006).

84 See generally Gillian E. Metzger, *Privatization As Delegation*, 103 COLUM. L. REV. 1367, 1400–10 (2003) (arguing that privatization would lessen judicial review); Steven L. Schooner, *Contractor Atrocities at Abu Ghraib: Compromised Accountability in a Streamlined, Outsourced Government*, 16 STAN. L. & POL'Y REV. 549, 562–64 (2005) (discussing the dangers of contractors who receive no guidance or management).

85 See *Hamdi v. Rumsfeld*, 542 U.S. 507, 518–19 (2004) (plurality opinion).

86 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).

review some decisions more stringently, and others less so.⁸⁷ 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.⁸⁸ 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 entrance to a national forest, and another (more stringent review) when courts are reviewing an intricate regulatory rule governing the licensing of nuclear reactors.⁸⁹

Stringency of review is what distinguishes these two different versions of the “arbitrary and capricious” standard. Stringency is 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. The concept includes 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 ought 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

87 See cases cited *supra* note 11.

88 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.D.2 for a discussion.

89 See, e.g., SECTION OF ADMIN. LAW AND REGULATORY PRACTICE, AM. BAR ASS’N, 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.

Id. (emphasis added).

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.⁹⁰ 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.⁹¹

Though courts engage in review at different levels of stringency depending on the statutory context and judgments about the feasibility of more intense scrutiny, their work across contexts tends to 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.⁹² In fact, courts treat horizontal equity as an important value, where deviations must be defended.⁹³ 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.⁹⁴ 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.

90 See *Hamdi v. Rumsfeld*, 316 F.3d 450, 473 (4th Cir. 2003), *vacated*, 542 U.S. 507 (2004).

91 See *Hamdi*, 542 U.S. at 509 (plurality opinion).

92 For some examples from different regulatory contexts, see *Bowen v. Mich. Acad. of Family Physicians*, 476 U.S. 667, 675 (1986) (discussing legislative reluctance to tax judicial resources by allowing routine review of Medicare Part B claims); *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”); *Langevin v. Chenango Court, Inc.*, 447 F.2d 296, 303 (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 assumes (either implicitly or explicitly) 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 12 for examples of briefs making this argument.

93 See *Cole*, *supra* note 2, at 2567–77.

94 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 the review of discretionary executive decisions is 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.

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 to prohibitively increase the associated costs in terms of the direct burdens of review and the forgone benefits of discretion.⁹⁵

D. Why Audits Could Serve as a Valuable Supplement to Traditional Judicial Review

How might the legal system better manage the costs and benefits of the high-discretion, low-review-stringency regime associated with actions committed to agency discretion or subject to highly deferential review? Plainly, the answer depends on how one defines "better," a point that will be taken up in earnest below. For the moment, imagine that society fears mistakes in the exercise of executive discretion but also seeks to harness expertise and conserve review resources. Imagine further that society has at its disposal some new amount of resources that can be used either to expand slightly the scope or stringency of traditional judicial review over highly discretionary activities, or in some other fashion. How else might those resources be used?

1. The Role of Audits in Overseeing Complex Organizations

One answer can be found in what government organizations repeatedly do to the public: they audit.⁹⁶ As the term is used here, an audit of executive discretion is a stringent evaluation of a sample of discrete decisions drawn randomly from a larger pool, using an explicit standard fixed in advance, with the results announced to the public.⁹⁷ Its aim is to uncover, for each reviewed case, whether a par-

95 Cf. CASS R. SUNSTEIN, *AFTER THE RIGHTS REVOLUTION* 219 (1990) ("[A]n unduly aggressive judicial posture may increase delays and paperwork in a way that threatens implementation.").

96 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.

97 Each of these features interacts to give the proposal developed here its unique characteristics. (1) Random selection assures a representative picture of decisions and provides a mixed strategy approach to deterrence that is difficult to evade. See *infra* Part II.A. (2) Fixing some standard in advance reduces risk-normalization dy-

ticular 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 program management audits, the audits of executive discretion envisioned here treat each discretionary decision, such as a decision to label a group as a specially designated terrorist organization, as the unit of analysis.⁹⁸ 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.”⁹⁹ 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.

Audits associated with taxation are among the most common. They take place in some form in most reliable tax collection systems.¹⁰⁰ 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.¹⁰¹ From this perspective, one of the “purest” tax auditing programs in recent years (in the sense of being

namics described by scholars of bureaucracy, making it more likely that problems will be detected by audit bureaucracies that are not immune from the organizational and political problems affecting the agencies they are auditing. See *infra* note 136 and accompanying text (discussing Vaughan’s work on the “normalization of risk”). (3) Public announcement (including provisions for protecting classified information) allows audits to play a role in a larger system of accountability defined by representative politics and public reactions. See *infra* Part II.D.

⁹⁸ Hence, the audits of executive discretion discussed here differ from the more informal program evaluation audits often undertaken by the GAO and Inspectors General. The incentives of federal audit bureaucracies to perform different types of audits are discussed *infra* Part IV.

⁹⁹ James G. March et al., *Learning From Samples of One or Fewer*, 2 *ORG. SCI.* 1, 2 (1991).

¹⁰⁰ See generally JOEL SLEMROD & JON BAKIJA, *TAXING OURSELVES* 180–82 (3d ed. 2004) (discussing how the IRS performs audits).

¹⁰¹ 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 (1990).

almost entirely random) was the Internal Revenue Service's Taxpayer Compliance Measurement Program (TCMP).¹⁰² 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.¹⁰³

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.¹⁰⁴

Audits also play a role in regulatory enforcement and court proceedings, such as when 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.¹⁰⁵ Courts reviewing this practice have repeatedly endorsed it, finding neither a conflict with the statute nor one with due process.¹⁰⁶ "Sample audits" also make an occasional appearance in class actions. When they do, courts confronted with an entire class of claims benefit from examining a sample of those claims to gauge the merits of the suit.¹⁰⁷

Private markets have made audits even more of a fixture. Administrators frequently deploy auditing techniques in the internal monitoring of private sector organizations, where the warning that "your call may be monitored for quality assurance" has become ubiquitous. Hierarchical organizations facing market pressures are riven with information problems. Managers may learn something about their subordinates' performance from market responses. But these

102 SLEMRD & BAKIJA, *supra* note 100, at 174.

103 *Id.* at 175.

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

105 *See* Chaves County Home Health Serv., Inc. v. Sullivan, 931 F.2d 914, 915 (D.C. Cir. 1991).

106 *Id.* at 916.

107 *See, e.g.,* McComber v. Travelers Prop. & Cas., 37 Conn. L. Rptr. 349, 350-51 (Conn. Super. Ct. 2004).

responses—commonly driven by a host of factors external to the organization—yield an inexact picture of performance, and waiting for market responses can risk the organization's well-being.¹⁰⁸ Audits are a response in a world of scarce resources. The point of such audits is not immediately to stop abuses or mistakes in discrete cases. It is to enable greater learning about what 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 of money paid out for a particular insurance claim was correctly calculated than they do paying out the claim itself.¹⁰⁹ 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.¹¹⁰

Organizational leaders value audits for a reason. Audits mitigate the problem, common to private organizations and 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. Recognizing the possibility of strategic action to evade monitoring, the Supreme Court lauded random enforcement in *United States v. Biswell*.¹¹¹ 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.¹¹² In holding that the Treasury agent could do so without a warrant, the majority observed that:

108 Cf. ALBERT O. HIRSCHMAN, EXIT, VOICE, AND LOYALTY 24–25 (1970) (discussing the risks to organizations of obtaining performance signals by simply waiting for market-driven "exit" responses among customers, employees, or other constituencies).

109 See J. David Cummins & Sharon Tennyson, *Controlling Automobile Insurance Costs*, J. ECON. PERSP., Spring 1992, at 95, 110–11. For a more general discussion of the value of, and incentives for use of, audits in the private sector, see Ross L. Watts & Jerold L. Zimmerman, *Agency Problems, Auditing, and the Theory of the Firm: Some Evidence*, 26 J.L. & ECON. 613, 626–33 (1983).

110 See KREPS, *supra* note 101, at 763–64.

111 406 U.S. 311 (1972).

112 *Id.* at 312.

[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.¹¹³

Although the Supreme Court was describing the needs of federal regulatory bureaucracies overseeing the public, it could just as well have been describing the needs of legislators overseeing the bureaucracy. Not surprisingly, legislators have created several bureaucracies capable of using auditing techniques to investigate what government agencies actually do with their discretion. Occasionally, government agencies audit the performance of their own workers.¹¹⁴ The Government Accountability Office (originally the General Accounting Office) was created early in the twentieth century primarily to help Congress monitor the financial activities of the executive branch.¹¹⁵ In 1974, legislators gave the GAO power to review and analyze the implementation of government programs.¹¹⁶ 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.¹¹⁷ Like the GAO, the Inspectors General have the legal power to investigate how federal officials use their targeted discretion. The existence of these structures indicates the potentially important role that audits can play in shaping how the federal government uses its targeted discretion. Whether these bureaucracies actually perform such audits is another matter, discussed below.¹¹⁸

How might audits be used in contexts where traditional review treats an executive decision as essentially committed to agency discretion, or subject to exceedingly deferential, low-stringency review? Imagine a world much like our own, where only some decisions are

113 *Id.* at 316.

114 See Exec. Order No. 12,564, 3 C.F.R. 224 (1987), reprinted in 5 U.S.C. § 7301 note (2000) (subjecting numerous categories of federal employees to random drug tests); Editorial, *Guilty Until Proven Innocent*, N.Y. TIMES, Jan. 20, 2006, at A16 (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 sixty-six percent of taxpayers deserved a full refund and another fourteen percent deserved a partial refund).

115 See John T. Rourke, *The GAO: An Evolving Role*, 38 PUB. ADMIN. REV. 453, 453-54 (1978).

116 *Id.* at 455.

117 See PAUL LIGHT, *MONITORING GOVERNMENT* 42 (1993).

118 See *infra* Part III.

subject to stringent judicial review, and others are subject to less stringent review. For fairly obvious reasons, political principals (such as legislators) 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 is 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.¹¹⁹

An audit of discretionary executive branch decisions would unfold in three steps. 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). The auditor would randomly choose some number or percentage of decisions to audit. Second, 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 IIEPA 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 so, the auditor would apply a standard either drawn in advance from the purposes of the statutes in question, or perhaps even based on the executive branch’s own assertions about the goals it seeks to

119 The proposal is, therefore, somewhat reminiscent of one that Mashaw offered for social security benefit determinations in the concluding pages of his exhaustive study of the Social Security Administration. See JERRY L. MASHAW, *BUREAUCRATIC JUSTICE* 226 (1983).

accomplish through the audited discretionary decision.¹²⁰ Intelligence information, reviewed in camera, could be used in these determinations.¹²¹ Third, 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.

Relying on public responses introduces some uncertainty about an audit's impact. 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.¹²² 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 other 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.¹²³

2. The Risks of Tolerating the Amount of Discretion Associated with Traditional Judicial Review

Why would such knowledge prove valuable? The first part of the answer will require us to retrace the most commonly asserted arguments for expansive discretion. By juxtaposing them against the risks of discretion, we can better identify the problems that arise from tolerating the amount of open-ended bureaucratic flexibility associated with traditional judicial review.

120 For instance, in the criminal context, the auditor's determination of a standard would be shaped by statements of prosecutors regarding the purposes of the prosecutions. See *infra* notes 204–05 and accompanying text.

121 Cf. *United States v. Isa*, 923 F.2d 1300, 1306 (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* 105–14 (1993) (discussing how fear of leaking is overblown).

122 See *infra* Part III.A.

123 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.

a. Arguments for Expansive Discretion and Their Contingent Nature

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 the speed with which a discretionary decision can be made. More review introduces delay in at least two ways: by providing incentives for executive branch decisionmakers to engage in greater analysis or deliberation before a decision is made (*ex ante*), and by potentially delaying the point at which a decision can be fully implemented until review is completed (*ex post*). Some decisions need to be made quickly if they are 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—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.¹²⁴ 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.¹²⁵ 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.¹²⁶

The argument for discretion in such cases often exalts expertise.¹²⁷ The conventional wisdom is that agencies and the executive

124 See *Water Keeper Alliance v. U.S. Dep't of Def.*, 271 F.3d 21, 30–35 (1st Cir. 2001).

125 See CORNELIUS KERWIN, *RULEMAKING* 174 (3d ed. 2003); see also McGarity, *supra* note 35, at 21–23 (discussing the effect of judicial review on agencies' willingness to use regulatory authority).

126 See *supra* Part I.C.

127 See *Perez v. FBI*, 71 F.3d 513, 514 n.6 (5th Cir. 1995). *But see* *N. Jersey Media Group, Inc. v. Ashcroft*, 308 F.3d 198, 219–20 (3d Cir. 2002).

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.¹²⁸ 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.¹²⁹ 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.¹³⁰

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. 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.¹³¹ 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

128 See *INS v. Lopez-Mendoza*, 468 U.S. 1032, 1048 (1984); *Vt. Yankee Nuclear Power Corp. v. Natural Res. Def. Council, Inc.*, 435 U.S. 519, 545-48 (1978).

129 Cf. KEITH KREHBIEL, *INFORMATION AND LEGISLATIVE ORGANIZATION* 265 (1991) (arguing that "informational concerns—in the sense of how politicians can be provided with incentives to study public problems and formulate public policy—are at the heart of legislative organization").

130 See *Greenbaum v. EPA*, 370 F.3d 527, 533-34 (6th Cir. 2004).

131 See, e.g., *Chevron v. Natural Res. Def. Council, Inc.*, 467 U.S. 837, 865 (1984) ("While agencies are not directly accountable to the people, the Chief Executive is . . ."); *Motor Vehicle Mfrs. Ass'n v. State Farm Mut. Auto. Ins. Co.*, 463 U.S. 29, 59 (1983) (Rehnquist, J., concurring in part and dissenting in part).

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."¹³²

b. The Risks Associated with Expansive Discretion

The arguments extolling discretion 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.¹³³ 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.¹³⁴ 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 that 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

132 *Newman v. United States*, 382 F.2d 479, 482 (D.C. Cir. 1967).

133 *See, e.g., MASHAW & HARFST, supra* note 35, at 111–23 (using the example of NHTSA to show that agencies react to judicial review).

134 *See* ARTHUR LUPA & MATHEW MCCUBBINS, *THE DEMOCRATIC DILEMMA* 101–48 (1998).

that does not go well.¹³⁵ Second, organizations develop routines that blind them. As Diane Vaughan wrote in her study of the *Challenger* launch decision:

Possibly the most significant lesson from the *Challenger* case is how environmental and organizational contingencies create prerational forces that shape worldview, normalizing signals of potential danger, resulting in mistakes with harmful 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.*¹³⁶

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.¹³⁷ 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.¹³⁸

135 This is certainly true in the case of people who work in offices whose broad performance is reviewed by the Inspectors General or the GAO. See *infra* Part III.A.

136 DIANE VAUGHAN, *THE CHALLENGER LAUNCH DECISION* 409–10 (1996) (emphasis added).

137 See March et al., *supra* note 99, at 2–3; Michal Tamuz, *Learning Disabilities for Regulators: The Perils of Organizational Learning in the Air Transportation Industry*, 33 *ADMIN. & SOC'Y* 276, 295–99 (2001).

138 See MASHAW & HARFST, *supra* note 35, at 111–23, 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: (1) that agencies appear to react to judicial review, and (2) that organizational learning is not necessarily associated with learning to 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

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,¹³⁹ 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 levers to create favorable public perceptions. Frozen assets and specially designated terrorist organizations send the message that the executive branch knows what it is 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,¹⁴⁰ diminished compliance with treaties, or simply the 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.¹⁴¹

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 have 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

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.

139 See Mariano-Florentino Cuéllar, *The Mismatch Between State Power and State Capacity in Transnational Law Enforcement*, 22 BERKELEY J. INT'L L. 15, 44–48 (2004).

140 See *id.* at 49–53.

141 See THE ABU GHRAIB INVESTIGATIONS 1–20 (Steven Strasser ed., 2004).

information available. Agency relationships change in response to what the players come to know, even if—to paraphrase former Secretary of Defense Donald Rumsfeld—voters know what they do not know.¹⁴²

A more plausible assumption can be grounded in the extensive behavioral research tradition in mid-to-late twentieth-century political science, suggesting that voters often do not know what they do not know.¹⁴³ 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.¹⁴⁴ 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,¹⁴⁵ but if they deliberately create them.¹⁴⁶ The point is not that electoral checks are irrelevant. It's that they are not entirely reliable, and they are often dependent on some mechanism—like judicial review—to focus public attention and produce information for it.

In effect, bureaucracies should frequently be expected to face pressures to render poor decisions, unless they are subjected to con-

142 See Hart Seely, *The Poetry of D.H. Rumsfeld*, SLATE, Apr. 2, 2003, <http://www.slate.com/id/2081042>.

143 See, e.g., LUPIA & McCUBBINS, *supra* note 134, at 17–38 (1998) (emphasizing citizens' limited knowledge about politics); Richard R. Lau & David P. Redlawsk, *Voting Correctly*, 91 AM. POL. SCI. REV. 585, 587–93 (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).

144 Cf. Mariano-Florentino Cuéllar, *Rethinking Regulatory Democracy*, 57 ADMIN. L. REV. 411, 412–16 (2005) (finding that comments made by citizens during the notice and comment phase of a new administrative regulation were less likely to be incorporated into the final regulation than comments made by administrative attorneys because the citizens' comments lacked technical sophistication even though they offered constructive insights relevant to the agencies' legal mandates).

145 See Kenneth E. Scott & Barry R. Weingast, *Banking Reform: Economic Propellants, Political Impediments*, in REFORMING FINANCIAL INSTITUTIONS AND MARKETS IN THE UNITED STATES 19, 27–36 (George G. Kaufman ed., 1994).

146 See Barry R. Weingast, *Caught in the Middle: The President, Congress, and the Political-Bureaucratic System*, in THE EXECUTIVE BRANCH 312, 338 (Joel D. Aberbach & Mark A. Peterson eds., 2005) (“Bureaucracy is a pejorative term in modern America, and rightly so. . . . The political compromises necessitated by the American constitutional system mean that legislation rarely attacks problems in a straightforward manner, but typically through political compromise that combines multiple and conflicting goals . . .”).

straints that nudge them in more 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. Under marginal cost analysis, if the costs of review only apply to a small fraction of cases, more stringent review is possible for 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.¹⁴⁷ In short, courts may find justification for avoiding intrusion into some domains of executive power on the basis of broad separation of powers or related political question rationales.¹⁴⁸ But the consequentialist case for limiting review of many routine executive branch decisions must contend with two realities that weaken the case for unfettered executive discretion: (1) the pervasive risks of bureaucratic mistakes or malfeasance, and (2) the fact that review mechanisms such as audits may provide valuable new information about executive performance without incurring the range of costs that traditional judicial review would.

3. A Social Welfare-Oriented Case for Auditing Executive Discretion

The preceding discussion about the risks of expansive discretion should tend to support the intuition that audits could improve on what traditional judicial review provides. We now turn to providing some analytical support for the intuition by considering a social welfare-oriented justification for audits in a politically uncomplicated

147 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 elaboration of the last point.

148 In any event, some observers suggest the political question doctrine itself is becoming increasingly irrelevant. For an insightful discussion of the doctrine's evolution, its relative decline, and the extent to which its theoretical bases are compelling, see Barkow, *supra* note 2. As Part I.B notes, no similar growth in judicial assertiveness is discernible in judicial supervision of more routine instances of executive discretion.

world, where traditional judicial review limits agents' ability to optimally police the exercise of executive discretion.

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 byproduct 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,¹⁴⁹ and that courts themselves occasionally make when they cautiously use samples of claims in a class, or health care reimbursement requests to learn more in the litigation process.¹⁵⁰

The argument for audits can be framed more explicitly by focusing on whether new resources available for reviewing executive discretion should be used to expand judicial review or to conduct audits of executive discretion. Recall that external review (from courts or otherwise) is subject to various types of resource constraints. Those constraints arise from the direct cost of review,¹⁵¹ the resources consumed by the executive branch in responding to review,¹⁵² and the risks of

149 See *infra* notes 190-91 and accompanying text (emphasizing the potential value, in social research design, of using scarce resources to generate more information about individual cases rather than increasing sample size).

150 See *supra* text accompanying notes 105-07 (discussing instances of sample adjudication).

151 See *supra* note 92 (referencing cases discussing the costs of review).

152 See *supra* note 12 (referencing federal government arguments discussing the dangers associated with consuming executive branch resources during the course of the review process).

over-detering desirable executive branch activity.¹⁵³ Given the scarcity of resources for review, imagine for simplicity that courts and legislators have used marginal cost analysis to optimally balance review stringency for all (actual or potential) cases arising in a particular class expected to arise at any point during a given time period.¹⁵⁴ And imagine further that additional resources (in the amount of R) become available to defray all of the aforementioned costs of review. How should they be deployed?

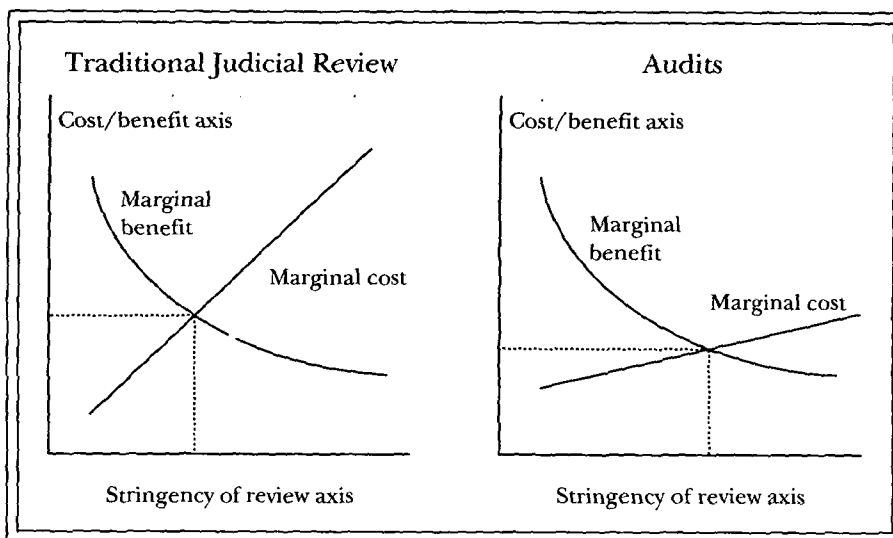
While no such question can be answered in a vacuum, some useful insights emerge if we make some additional (stylized but plausible) assumptions. First, audits—as the term is used here—involve random selection of relevant discretionary decisions for review from a total pool of decisions made in a given time period.¹⁵⁵ Moreover, the expected costs of review of all eligible cases under traditional judicial review in a given time period are higher than for audits. In marginal cost and benefit terms, as Figure 1 shows, with audits the costs of review rise more slowly as stringency rises because they apply to only a fraction of the total pool of (actual or potential) cases litigated. This property of audits allows the optimal degree of stringency to be higher for cases audited than for cases subject to traditional judicial review.

153 Cf. PETER H. SCHUCK, *SUING GOVERNMENT* 59–81 (1983) (analyzing the over-deterrence problem in the context of damage actions against the government).

154 Court opinions often include references to the efficiency of the particular degree of stringency they have chosen (or the legislature has) for a given class of cases. See *supra* note 92. But courts may fail at conducting the analysis. They may be concerned about workload pressures to a degree that overwhelms more nuanced evaluation of marginal costs and benefits. Cf. Theodore Eisenberg & Stewart Schwab, *The Reality of Constitutional Tort Litigation*, 72 *CORNELL L. REV.* 641, 693–95 (1987) (discussing the disjuncture between judicial workload concerns and the actual extent of litigation involving constitutional torts). They may fail to analyze the full range of benefits and costs associated with more (or less) stringent review. Cf. Cass R. Sunstein & Adrian Vermeule, *Interpretation and Institutions*, 101 *MICH. L. REV.* 885, 890 (2003) (arguing that theories of interpretation that ignore institutional capacities and the dynamic effects of any particular approach to interpretation are inadequate). In any case, the argument for audits developed above is strengthened when it turns out that courts and legislatures fail at balancing marginal costs and benefits when fixing the level of stringency for a particular class of cases.

155 For a discussion of sample size, see *infra* Part II.B.

FIGURE 1. OPTIMAL STRINGENCY UNDER AUDITS AND TRADITIONAL REVIEW

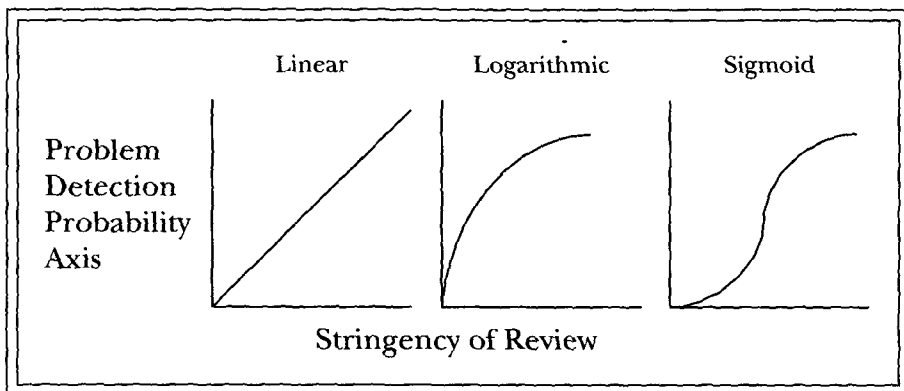


Second, for any individual case reviewed, the probability of discovering problems increases as a function of stringency of review. That is, the harder a reviewer looks, the greater the chance she will discover a problem with a discretionary decision if there is one.¹⁵⁶ But there is an additional complexity, arising from variations in the shape of the function representing the relationship between review stringency and problem detection. The importance of this function should be reflected in any principled analysis of the marginal benefits of review. As Figure 2 indicates, the stringency-problem detection function may be linear, logarithmic, or sigmoid. Variations presumably depend on context: willful malfeasance should tend to be associated with a sigmoid (s-shaped) function, because of perpetrators' efforts to cover their tracks, thereby making it possible to discover problems only after considerable stringency is used to review a decision. For decisions having a sigmoid stringency-problem detection probability curve, the probability of discovering problems under highly deferential review may be essentially zero—a condition that is likely to change only with relatively large increases in review stringency.

¹⁵⁶ For a discussion of what standard a reviewer would apply in an environment where discretion is often defined precisely by the *absence* of standards, see *infra* Part II.C.

Third, there is a distinctive social benefit (call it an “adjudicatory” benefit) associated with the use of resources for traditional judicial review (as opposed to audits) because of the special role that process plays in resolving individual disputes in addition to generating information.¹⁵⁷ The benefit varies depending on whether the new review resources are likely to correct instances where litigants have suffered without justification as a result of an executive branch discretionary decision.

FIGURE 2. DIFFERENT STRINGENCY- PROBLEM DETECTION CURVES



Fourth, we assume there is relative social consensus regarding the broad goals associated with lodging discretion for the type of decision in question within the executive bureaucracy. Difficulties arise because achieving those goals requires both the detection of errors by bureaucracies and flexibility allowing executive branch officials to apply expertise in specific circumstances. Nonetheless, the extent of consensus regarding overall goals in this context is reflected in a stable majoritarian coalition of legislators and relevant executive branch officials (though neither the goals in question nor particular interpretations of what states of the world achieve those goals are necessarily shared by everyone). We relax this assumption below.

Because we have assumed only R new resources, the specific question is when it makes sense to allocate these resources to audits rather than traditional judicial review, and assume for simplicity that this is an all or nothing, binary choice (though nothing critical turns on this). In an ideal world, that choice ought to depend on several factors that are important to optimize limited review resources. One crit-

¹⁵⁷ This property also generates its own costs, and litigants may face a variety of standing, redressability, and remedy-related hurdles if they choose to litigate their cases. But for simplicity, these details need not be addressed here.

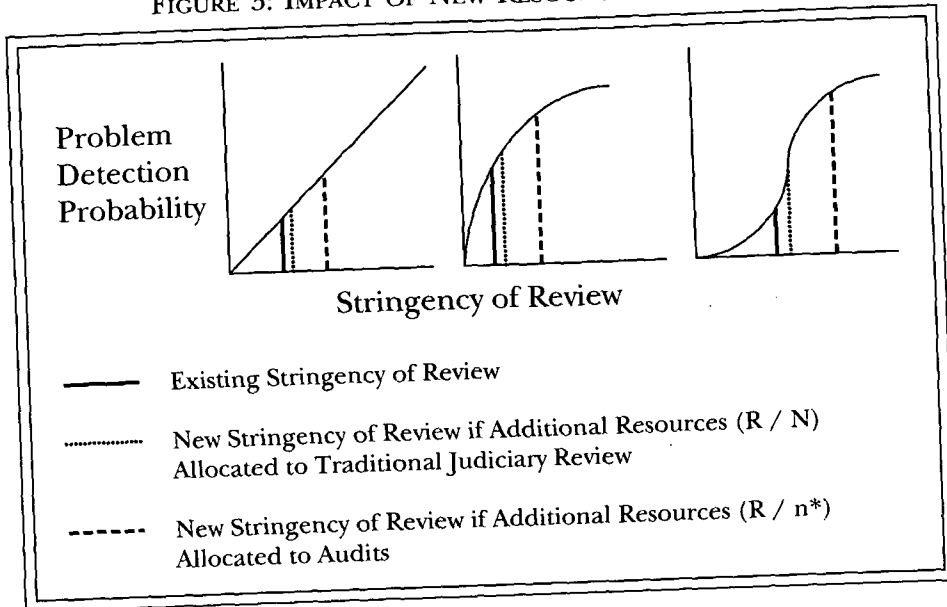
ical factor is the type of stringency-problem detection curve most common in a particular context, and where the status quo fits in relation to that curve. Suppose again we are dealing with a sigmoid curve. As Figure 3 shows, if the optimal stringency of review under traditional judicial review before the new resources are added is associated with a low point on the problem-detection axis, then the new resources are likely to do very little for traditional judicial review. The available resources must be divided across all cases in a particular class, making it unlikely that the problem-detection probability will rise much. Other critical factors include the value of additional information generated from reviewing all cases instead of a sample of them, and the "adjudicatory" benefit that might arise if resources spent on adjudication materially change the probability that litigants will be able to address problems associated with discretionary decisions that have aggrieved them.

In a world with the aforementioned characteristics, the case for spending the resources on traditional judicial review instead of audits turns out to be a difficult one. As is almost certainly the case in the real world, we have assumed that stringency-problem detection curves vary depending on the context. Only rarely will they be of the precise shape necessary to render pivotal the small (if not tiny) increase in review stringency at the margin that is possible when dividing R by the total number of cases eligible for review under traditional judicial review (N). In contrast, audits can generate valuable increases in problem detection probabilities by allocating R among a smaller pool of sampled cases (n^* in Figure 3). Perhaps something valuable is gained from reviewing the full population of cases instead of a sample? Counterintuitively, sampling theory suggests that—from an information-generation perspective—the marginal increase in the accuracy of information available by reviewing more cases declines dramatically after a minimum number of cases (several hundred, for example) are selected for review at random, and this decline occurs even when the total number of cases in a population is vast.¹⁵⁸ Moreover, unless R is relatively large compared to the previously available resources for review, the value of using the new resources for traditional review from a

158 See *infra* Part II.B for details. Sampling theory provides one important rationale for randomization—as opposed to an intuitive but potentially misleading focus on “problem” cases. Although audits can be adapted to focus on reviewing a subset of decisions that seem from surface indicators to be especially likely to be problematic, such an approach dilutes what can be learned about the entire pool of cases. To the extent the relevant bureaucracies and associated organizational leaders know the basis for selection, it also makes it easier to evade review. See *supra* note 101 and accompanying text (discussing mixed strategies).

litigant-focused perspective proves questionable. Unless the stringency-problem detection curve turns out to make that tiny sliver of resources available per case critical, there will be little change in review practices from a litigant's perspective.¹⁵⁹

FIGURE 3: IMPACT OF NEW RESOURCES ON REVIEW



But there is still another problem with using the resources for traditional judicial review. Suppose a particular class of decisions subject to review—such as EPA administrative compliance orders, or Department of Homeland Security (DHS) “no-fly” list designations—featured the precise stringency-problem detection curves, making them highly sensitive to tiny increases in stringency. The value of traditional judicial review as a tool for modern bureaucratic performance would be further limited by the information-distorting properties of litigation. 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.¹⁶⁰ As William Simon has pointed out in the wel-

¹⁵⁹ And suppose the right stringency-problem detection curve were present—making the new resources critical for each case reviewed. Such a scenario would increase the desirability of litigation, further increasing the costs associated with traditional review.

¹⁶⁰ There may be both overbreadth and under-inclusiveness 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, 1–6 (1984). For an illuminating

fare 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).¹⁶¹ In contrast, audits have the potential to generate information regardless of the willingness or opportunity of individuals to challenge specific decisions.

Although the preceding account abstracts in limited ways from the real world, it helps demonstrate why audits would be valuable even in a world already providing some review of discretion. In most cases where such review is absent or otherwise undermined, the value of audits would be even greater. The prescriptive case for audits potentially holds in more complex specifications, such as where political goals are controversial or resources are redistributed from traditional judicial review and plowed into audits. But the case is especially compelling in the present analysis, because the question involves how to use new resources, and there is a broad social consensus about what executive bureaucracies should accomplish.

Evidently, audits introduce their own costs into the equation. While fixing the precise cost of an audit system depends on the institutional design issues taken up in Part II, 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. If 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. Presumably, the costs associated with audits, compared to extending high-stringency judicial review to all cases in a particular class, would be less—even if we kept the potential remedy—than reviewing all potential cases. Recall: audits would not necessarily yield a direct remedy.¹⁶² To the extent that analyses of the marginal costs and benefits of review incorporate the possible costs to the executive branch of

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, 78–85 (2004).

161 See William H. Simon, *Rights and Redistribution in the Welfare System*, 38 *STAN. L. REV.* 1431, 1467–68 (1986).

162 If determined advocates of executive branch discretion choose to treat the potential loss of public trust arising from audits 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.

having a decision vacated by a court, audits would also involve lower costs because they can be designed merely to reveal information rather than to provide direct relief.¹⁶³

Although 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 have 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.¹⁶⁴

Second, the belief that audits can deliver the aforementioned benefits depends on making certain assumptions about the powers of the auditor, though all are plausible. The auditor must 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: (1) that the auditor is

163 For a brief discussion of whether the GAO and Inspectors General should allocate a material share of their existing resources to audits of executive discretion, see *infra* Part III.A.

164 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 a broader arbitrariness review). In those exceptional circumstances, the framework I have provided suggests that sometimes—such as when it is 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.

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; (2) 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 (3) 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 is 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.¹⁶⁵ Nonetheless, the auditor(s) must know something about what constitutes accuracy when allegedly terrorist assets are frozen or when agencies use administrative compliance orders.¹⁶⁶ 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.

4. How Audits Would Fare in a Politically Fragmented World

Most of the time, the world in which bureaucracies exercise discretion—and therefore the world in which auditors would operate—is more complicated than the analysis above suggests. Politicians differ over what they want the bureaucracy to accomplish.¹⁶⁷ Even when leg-

165 See, e.g., VAUGHAN, *supra* note 136, at 63 (“[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) (“A warning . . . is only effective if it fits into our mental model of what is going on. As with the ‘warnings’ of Pearl Harbor, it can get swamped by the multitude of signals that fit our expectations, and thus be discounted as ‘noise’ in the system.”); SCOTT D. SAGAN, *THE LIMITS OF SAFETY* (1995) (noting that adding redundant safety devices does not always increase safety because human operators can often work around these redundancies if the devices do not serve immediate interests).

166 And obviously, whatever the benefits of specialization, they need to be balanced against the risks.

167 See generally Terry M. Moe, *The New Economics of Organization*, 28 AM. J. POL. SCI. 739 (1984) (arguing that theories of hierarchy developed in the field of organizational economics can contribute much to the understanding of political control over bureaucracies).

islative majorities support the creation of the bureaucracy or the delegation of legal powers to it, some coalition members may actively work to restrain the agency's powers. Others will encourage the agency to make the broadest reading of its mandate.¹⁶⁸ In what follows, I relax the assumption of broad political consensus about what constitutes desirable bureaucratic performance. The analysis reveals: (1) that auditors would likely have a much harder time in such a world, but (2) that participants in the political process could still be persuaded to support audits in narrow circumstances. Put simply, the analysis here shows how generating sufficient political interest for audits to survive is possible even in a politically complicated world of heterogeneous political preferences and divided politicians. The range of dynamics making this result unlikely (though far from impossible) is explored in Parts III and IV. For now the point is that political realities make the fate of audits *contingent*, rather than dooming them entirely.

To see this, imagine a somewhat different situation than the one posited in the previous subsection. First, lawmakers and executive branch officials in a governing coalition have political preferences regarding what they want the bureaucracy to do with its legal authority (approve drugs faster or more slowly, impose fewer or more environmental administrative compliance orders, make it easier or more difficult for an alien to demonstrate sufficient credible fear to merit formal asylum procedures, and so on).¹⁶⁹ And those preferences differ more often than not. Second, politicians in a governing coalition may nonetheless *also* have a set of more widely shared concerns about bureaucratic performance.¹⁷⁰ Regardless of their preferences about the existence of nuclear weapons, for example, lawmakers and executive branch officials are likely to be alarmed by material lapses in the

168 See Weingast, *supra* note 146, at 318–31.

169 For examples of work analyzing the implications of divergent preferences on specific issues among members of broad governing coalitions, see David P. Baron, *A Noncooperative Theory of Legislative Coalitions*, 33 AM. J. POL. SCI. 1048 (1989); Michael Laver & Kenneth A. Shepsle, *Government Coalitions and Intraparty Politics*, 20 BRIT. J. POL. SCI. 489 (1990). For an analysis of when coalition governments have especially pronounced incentives to monitor the bureaucratic implementation of policies at the ministerial level, see Micahel F. Thies, *Keeping Tabs on Partners: The Logic of Delegation in Coalition Governments*, 45 AM. J. POL. SCI. 580 (2001).

170 See Bruce Bimber, *Information as a Factor in Congressional Politics*, 16 LEGIS. STUD. Q. 585, 596 (1991). Indeed, although the incidence of legislative proposals drawing unanimous or near-unanimous support varies, such proposals are not rare. For a discussion of empirical evidence demonstrating the recurring importance of universal or near-universal legislative coalitions, see Melissa P. Collie, *The Legislature and Distributive Policy Making in Formal Perspective*, 13 LEGIS. STUD. Q. 427, 445–49 (1988).

safety of a nuclear stockpile.¹⁷¹ Concern about bureaucratic malfeasance may also straddle ideological divisions even when the potential for catastrophe falls far short of accidental nuclear detonation; both opponents and proponents of particular surveillance programs, for instance, may share concerns about the program's potential mismanagement even when their underlying reasons for concern are different.¹⁷² Third, audits consume resources. As the analysis above noted, these resources may be used for alternative purposes. The political merits of those alternative uses weigh on decisionmakers. Fourth, politicians may have greater or lesser uncertainty about bureaucratic performance depending on the context.¹⁷³

When politicians consider the merits of developing an auditing system in such a world, the complications come from more than just the direct cost of that system. While audits may limit near-universally derided problems (such as, for example, federal agents' use of surveillance authority under the Patriot Act to pursue personal vendettas), they can also stop agencies from achieving goals that some politicians would rather see them achieve.¹⁷⁴ Politicians consider this trade-off in the context of four critical factors that follow from the assumptions above: (1) the cost of audits relative to other desirable functions, (2) the perceived risk of broadly derided problems, (3) the extent of uncertainty about bureaucratic performance with respect to such problems, and (4) the potential impact of audits on agency activities that inspire differing degrees of political support. It is manifestly possible, as lawmakers and executive branch officials weigh these considerations, for them to conclude that audits' potential to reduce uncertainty and prevent near-universally derided bureaucratic activity outweigh their costs or their unfavorable political effects.¹⁷⁵

171 For an illuminating account of how common such lapses are, see SAGAN, *supra* note 165, at 251.

172 See *supra* Part I.D.2 for a discussion of the risks of bureaucratic malfeasance and related shortcomings.

173 See Jonathan Bendor & Terry M. Moe, *An Adaptive Model of Bureaucratic Politics*, 79 AM. POL. SCI. REV. 755, 761 (1985) (discussing variations in the extent of politicians' knowledge of bureaucratic performance).

174 Audits' chilling effect on agency activities found desirable by a subset of the political coalition would arise as a result of the information they might reveal to pivotal legislators or organized interests, who could in turn shape the perceptions of mass political audiences.

175 Cf. William Bernhard, *A Political Explanation of Variations in Central Bank Independence*, 92 AM. POL. SCI. REV. 311, 312-15 (1998) (discussing how legislators may create an independent central bank as a commitment device not only in spite of, but because of, politicians' differing incentives on monetary policy); Matthew Potoski, *Managing Uncertainty Through Bureaucratic Design: Administrative Procedures and State Air*

Such calculations would be shaped by beliefs about the prospects for political control of the bureaucracy, which can raise or lower the value of keeping bureaucratic discretion as unfettered as possible from their perspective.¹⁷⁶ In addition, as noted in Part IV, politicians who might support audits if voting sincerely may still have strategic reasons to oppose audits in the hopes of provoking more pronounced institutional change. But the possibility remains that audits initially can survive political evisceration, and the intuition in support of that possibility is a simple one: even in the presence of ideological differences, raw errors can matter enough to engender support for audits among politicians and executive branch officials with sharply different policy goals.¹⁷⁷ And once audits begin to occur, the auditors themselves may endeavor to build a political constituency for them.¹⁷⁸

The strategic dynamic occasionally making audits politically viable can be analogized to committee deference in legislatures, where majoritarian chamber coalitions defer to legislative committees to encourage the development of specialization. The problem legislators face in that context is uncertainty about the precise effect of the complex laws they enact, such as immigration and refugee law provisions, complex tax depreciation changes, new rules governing the transfer of defense-related technologies to other countries, and changes to the drug approval process. An important strand of scholarship on the organization of the Congress persuasively argues that politicians can better achieve their goals by reducing the extent of that uncertainty, even when lawmakers have competing or opposite distributional inter-

Pollution Control Agencies, 9 J. PUB. ADMIN. RES. & THEORY 623, 626 (1999) (discussing the information-generating benefits of specific procedural mechanisms and institutional designs for state legislators with differing political goals); Douglas H. Shumavon, *Policy Impact of the 1974 Congressional Budget Act*, 41 PUB. ADMIN. REV. 339, 346–47 (1981) (discussing legislators' incentive for the creation of the Congressional Budget Office on the basis of the value of unbiased information to politicians with differing political goals).

176 See Rui J.P. de Figueiredo, Jr., *Electoral Competition, Political Uncertainty, and Policy Insulation*, 96 AM. POL. SCI. REV. 321, 362–28 (2002) (demonstrating the extent to which, from legislators' perspective, the desirability of mechanisms to constrain the executive branch depends on beliefs about future control of the executive branch).

177 The existence of an audit program—run by an external private-sector company—associated with the Treasury Department's anti-terrorism wire-transfer screening program involving the SWIFT system may serve as an example. See Greg Miller & Josh Meyer, *Officials Defend Bank Data Tracking*, L.A. TIMES, June 24, 2006, at A1. Some politicians opposed to the program would still prefer that, if the program is to operate, audits be included in its institutional design. See, e.g., *id.* But see *infra* Part IV.A for a discussion of why some opponents of the wire-transfer screening program might have strategic reasons to oppose audits.

178 See *infra* Part IV.C.

ests.¹⁷⁹ On the other hand, reducing uncertainty is costly and requires specialization. Politicians solve the problem by organizing the legislature to encourage specialization, thereby reducing uncertainty over time and enlarging the scope of benefits legislators can distribute. As with the opportunity cost associated with specialization, the creation of and deference to an audit system introduces a layer of costs. Some costs—such as allocating the resources for the system to operate—are generic, borne across the legislature itself. Others, such as the potential impact of audits on the ability of bureaucracies to make discretionary decisions stretching their legal mandate and pleasing to only a fraction of a governing political coalition, are borne by specific members (and in this case, those preferring such potentially distorted implementation). But given the limitations of traditional judicial review and other monitoring strategies, *not* having audits probably increases legislators' uncertainty about the consequences of their policy choices, raising the same concerns for legislators that the absence of specialization would.

There is, in effect, a limited political window for audits that opens when political fears of bureaucratic mistakes and malfeasance combine with politicians' uncertainty about bureaucratic performance. A politically-oriented justification for audits is possible even in a complicated world of politicians with competing agendas, where those agendas may include a desire among some legislators and executive branch officials to see bureaucracies fail in achieving their stated goals. Since neither political apathy nor political sabotage are guaranteed to doom randomized audits, we may turn our attention to the intricacies of audits' potential institutional design.

II. WHY THE INSTITUTIONAL DESIGN PROBLEMS INHERENT IN AUDITS ARE MANAGEABLE

While a framework for conducting audits is adaptable almost by definition, the design of that framework depends on the answers to practical questions. Four questions, in particular, prove central to implementing audits. Rather than seeking to resolve all the intricacies of an audit system, the discussion that follows instead shows how to address the four critical institutional design questions and how the issues raised by those questions can be dealt with in practice.

179 See KREHBIEL, *supra* note 129, at 264–65; see also Thomas W. Gilligan & Keith Krehbiel, *Organization of Informative Committees by a Rational Legislature*, 34 AM. J. POL. SCI. 531, 546 (1990) (noting that committees are superfluous if they possess no expertise with which to reduce uncertainty regarding the effects of legislation that they sponsor).

A. *What Executive Decisions Should be Audited?*

The most immediate such question concerns what discretionary decisions should be prioritized for auditing in a world of scarce resources. 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.¹⁸⁰ But which decisions should be prioritized for audits? It is plainly obtuse to seek audits of the most trivial discretionary actions and informal adjudications, or for similarly trifling executive decisions outside the context of federal administrative law.¹⁸¹ Certain routine discretionary decisions, showcased in the preceding discussion, tend to involve the application of legal authority or general policy justifications to specific facts; their primary effect is on specific individuals and groups.¹⁸² This type of discretionary power is what allows executive branch officials to freeze allegedly terrorist assets, place someone on a government-run “no-fly” list, or levy certain environmental or occupational safety fines. This 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 (e.g., “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.¹⁸³ Because of their frequency and their impact on discrete individuals and groups, these “application to the fact” types of routine discretionary decisions may be most immediately suitable for audits.

180 DAVIS, *supra* note 2, at 216.

181 Cf. 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, 1069 (2004) (“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.’”).

182 Routine executive discretion involving the targeted application of standards to individuals or groups 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 *id.* at 1069–75. 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 of 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.

183 See Brief for the Respondents, *supra* note 12, at 11–12.

Of course, not all instances of routine discretion have the preceding characteristics. Within the category of routine discretion outside the realm of what would be protected by the political question doctrine, some observers might legitimately distinguish decisions involving broad policy judgments. Certain discretionary judgments, for instance, 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. Although the principle associated with audits (trading off breadth in exchange for greater depth) could also be usefully applied to instances of broad policy judgments, broadly applicable policy judgments may call for a different form of supplemental review compared to the routine discretionary decisions that involve applying rules or standards to a particular set of facts.¹⁸⁴

No doubt sometimes executive branch officials will argue that the distinction between broad policy judgments and routine applications of authority to individuals and groups is illusory, because 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 alleged terrorist assets, the State and Treasury Departments implement statutory standards from the International Emergency Economic Powers Act (IEEPA),¹⁸⁵ as interpreted through policy judgments in the President's executive orders. They may further contend that the details of a policy judgment—like precisely what behaviors make a charity liable to have its assets frozen—

184 Moreover, such judgments are more likely to garner media and legislative attention even in the absence of audits. And existing investigations performed by audit bureaucracies and congressional committees are more likely to analyze broad policy judgments rather than routine instances of executive discretion applying standards to individuals or groups. See *infra* Part III. Hence, when resource constraints force a choice between using audits to screen instances of routine discretion involving applications of law to the facts and broader policy judgments, the former probably merits greater attention. More generally, the arguments for deferring to the executive branch when an agency writes a rule or a president signs an executive order seem 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."

185 50 U.S.C. §§ 1701–1706 (2000).

should not be made entirely public,¹⁸⁶ or that 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.¹⁸⁷

Even if one finds these positions attractive on the surface, it remains useful to draw *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 discretion in accordance with a particular policy rationale or standard (even one that is not derived directly from a statute), arguments against review become exceedingly difficult to accept. Audits would generate information about whether the publicly asserted standard was reflected in the discretionary decisions. Though some may insist that certain discretionary decisions involving national security (for example) are entirely immune from any standard,¹⁸⁸ 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.¹⁸⁹

186 See *Doherty v. U.S. Dep't of Justice*, 775 F.2d 49, 52 (2d Cir. 1985). For a skeptical view of such contentions, see VAUGHAN, *supra* note 136, at 389–90.

187 Cf. *Marbury v. Madison*, 5 U.S. (1 Cranch) 137, 170 (1803) (explaining that although the Court could not review decisions of the executive branch that involved political discretion, it could enforce the rights of individuals who had been harmed by executive actions that did not involve political discretion).

188 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 is too difficult to monitor, 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.

189 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, 1212 (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—which

Because of scarce resources, the exclusion of broad policy judgments from attention would still force auditors to make additional trade-offs in choosing where to focus their attention. In doing so, auditors could 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.

B. What Sample Should Be Used?

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.¹⁹⁰ 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 structur-

they now discourage, though audits might be a promising alternative to simply banning such rules—instances of information dissemination and use of regulatory agencies (as an alternative to the potentially cumbersome data quality act rules).

190 See SAMPLE SIZE CALCULATION IN CLINICAL RESEARCH 28–29 (Shein-Chung Chow et al. eds., 2003).

ing 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, 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.

Those choices regarding sample size should also reflect the fact that sample size has diminishing marginal returns. Counterintuitively, 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 500 and 1000 observations may provide a revealing statistical snapshot even when drawn from exceedingly large populations.¹⁹¹ 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.¹⁹²

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

191 For example, if a sample of 1000 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 5000 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 & MATTHEW DAVID, *SOCIAL RESEARCH* 154 (2004).

192 An exhaustive analysis of twenty-five 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 191, at 69.

(for example, the total number of indictments issued over several years, or the total number of disability determinations).¹⁹³ 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.¹⁹⁴ Finally, regardless of the size of the population, the 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 is 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. The first problem 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 National Security Council but, perhaps because of political considerations, were not specially designated). The former is more accurate

193 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 Determination in an Audit of Property Value Appraisals*, REAL. EST. ECON., Spring 2006, at 133.

194 Cf. SUTTON & DAVID, *supra* note 191, at 154 (noting that even a sample as small as thirty 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 99, at 2 (noting that sometimes “[o]rganizations attempt such pooling” of cases to create a larger sample from which to assess performance).

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 expanded 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 could have the laudatory effect of giving auditors reason to cultivate reputations for impartiality.

C. *Which Organization Should Audit and What Standard
Should It Apply?*

The prospect of overcoming political resistance to audits depends in part on what institutional actor shoulders the burden of auditing. Despite the existence of cases involving sample audits, some observers may question whether ordinary courts would have the inclination, legal authority, institutional culture, or expertise necessary to engage directly in audits.¹⁹⁵ The prospect of court involvement in auditing may be simplified if judges appoint masters to undertake 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.¹⁹⁶ 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 una-

195 For a somewhat stylized argument that the Supreme Court (or, more generally, appellate courts) *already* engage in something akin to audits, see Matt Spitzer & Eric Talley, *Judicial Auditing*, 29 J. LEGAL STUD. 649 (2000). It should be noted, however, that Spitzer and Talley are using the concept of auditing differently in that context, to describe a court's decision to review a particular case at a high level of stringency rather than a decision to select a subset of executive branch discretionary decisions at random for more stringent review. See *id.* at 652.

196 They would have to do this in a way that avoids running afoul of the Supreme Court's decision in *Vermont Yankee Nuclear Power Corp. v. Natural Resources Defense Council, Inc.*, 435 U.S. 519 (1978).

ble to require audits given constraints in their ability to impose procedures not explicitly grounded in statutes,¹⁹⁷ legislators could create an Article I court with a distinctive mission and resources to build specialized capacity—or an entirely separate bureaucracy. Among existing agencies, the GAO and IG Offices are best positioned to do this sort of work (though, as I note below, they have largely avoided doing so). In short, 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. At the same time, direct participation by courts is not essential to the success of an audit scheme,¹⁹⁸ the results of which would permit courts to vary stringency of review for different classes of decisions even if judges are uninvolved in conducting the audits themselves.

Dispensing with courts as auditors raises the question of whether *any* government agency is necessary at all. References to random selection and evaluation standards set in advance admittedly turn the argument for audits into what resembles a plea for more research, with the nature of the researcher mattering less than the contents and methods of the analysis. But the situation is not quite so simple, and as a prescriptive matter, cuts in favor of audits by a public agency. Government bureaucracies have powers and responsibilities that researchers rarely do. Independent regulatory agencies showcase how public agencies can be transparently engineered for balance.¹⁹⁹ Agencies may be able to solve collective action problems by deploying subpoena powers, and more readily generating media attention. The arguments for government conducted audits therefore echo some of

197 *Id.* at 557–58.

198 Audits would not deprive courts of alternative means of adjusting stringency of review even within a particular class of cases. Spitzer and Talley, for example, discuss how reviewing courts could combine their beliefs about a lower court and observable characteristics of a lower court's decision to determine whether to review that decision (or to review it more stringently). Spitzer & Talley, *supra* note 195, at 667 (indicating how appellate courts, functioning as “auditors” of lower courts, might adopt “an auditing strategy that can be either even-handed or asymmetric, depending on whether the dominant reason for reviewing is (respectively) imprecision or ideology”). The authors recognize, however, that such an approach represents a departure from the principle that like cases should be treated alike by courts, and acknowledge that external observers expect courts to adhere to that principle. *See id.* at 675 (“One of the basic principles of Anglo-American jurisprudence is that like cases should be treated alike.”).

199 *See* John F. Duffy, *The Death of the Independent Regulatory Commission (and the Birth of a New Independence?)*, 54–57 (June 9, 2006) (on file with author).

those for government regulation of food and drugs, for example, instead of self-regulation.²⁰⁰

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.²⁰¹ 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.²⁰² In addition to courts reviewing the information in camera, high profile commissions like the 9/11 Commission and expert working groups routinely get security clearances and access to classified information.²⁰³ 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

200 Cf. SUNSTEIN, *supra* note 95, at 45 ("In some circumstances, [government] interference may produce significant benefits for trivial costs. . . . [R]egulatory programs are necessary to solve collective action and coordination problems.").

201 Under existing law, judges sometimes handle sensitive information and use it to make decisions in cases. See *United States v. Nixon*, 418 U.S. 683, 715–16 (1974). It is 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 is 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, 24–25 (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).

202 See, e.g., *Jifry v. FAA*, 370 F.3d 1174 (D.C. Cir. 2004).

203 See NAT'L COMM'N ON TERRORIST ATTACKS UPON THE U.S., *THE 9/11 COMMISSION REPORT* (2004).

discretionary powers.²⁰⁴ The auditor could even use statements from the executive branch itself to see whether the audited cases seem to be consistent with those statements.²⁰⁵ 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 contexts, when they audit “broad management practices”). The standard might reflect insights drawn from constitutional interpretation, policy considerations, or even statutes’ legislative history.

D. *What Should Depend on the Outcome of an Audit?*

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 de-

204 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. § 701(a)(2) (2000). *Chaney*, 470 U.S. at 837–38. 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, 192–94 (1993). What these cases sometimes gloss over (but seem to recognize far more explicitly in cases involving the non-delegation 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, 1162 (D.C. Cir. 1973) (finding an exception where an agency has “consciously and expressly adopted a general policy” that is so extreme as to amount to an abdication of its statutory responsibility). This is not to suggest that courts should be the primary actors in conducting targeted discretion audits (perhaps some independent auditing authority including a mixture of judges and non-judges would have more flexibility to articulate standards). It is, rather, to point out that courts’ reluctance to articulate standards when they find them missing on the face of a statute should not be taken as an indication that such an enterprise is fruitless.

205 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 is plausible and announce the results, or it could rely on agency definitions of “national security” in related contexts. Cf. *Chaney*, 470 U.S. at 836 (suggesting that an agency decision not to enforce could be challenged if the agency itself has committed to act in specified circumstances).

depends on the reactions of legislators and the mass public, which can vary depending on the circumstances.²⁰⁶ 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 have generated dozens of stories in national newspapers and television networks.²⁰⁷ Thus, public disclosure may have the potential to impact crucial features of the political game as the executive branch seeks to demonstrate its competence.

On the other hand, mere public discourse may accomplish little where the media, organized interests, and the mass public have other concerns. Recall that 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.²⁰⁸ 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 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 the politi-

206 See Asher Arian & Sigalit Olzaeker, *Political and Economic Interactions with National Security Opinion: The Gulf War Period in Israel*, 43 J. CONFLICT RESOL. 58, 65-71 (1999); David O. Sears et al., *Political System Support and Public Response to the Energy Crisis*, 22 AM. J. POL. SCI. 56, 66-75 (1978).

207 See *infra* notes 225-28 and accompanying text.

208 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.

cal considerations discussed in Part IV. The key is to evaluate the problem of remedies 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 their 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 CONTRAST WITH THE STATUS QUO

The analysis began by focusing primarily on court review of executive discretion. In reality, though courts have distinctive legal powers and responsibilities, they are but one component of a larger web of institutions potentially capable of overseeing executive discretion. Formal judicial review is but one means through which executive power may be restrained—and even this depends to a considerable extent on its interaction with the views of government officials, organized interests, and the public regarding the value of compliance with judicial decisions.²⁰⁹ Other factors include some combination of the press, audit bureaucracies such as the GAO and Inspectors General, legislators, and even organized interests and social movements drawing support from among the mass public.

While these factors are unquestionably relevant, two observations are worth making about their impact on executive discretion. First, in many cases, legislators, organized interests, the media, and the mass public become a viable means of disciplining executive discretion *in response* to the information generated through formal review mechanisms, such as the judiciary's role in overseeing executive activities.²¹⁰ Because of judicial review's importance, limitations on judicial review may distort the picture of executive discretion that emerges—and therefore the relevance of the aforementioned factors. Politicians and the media sometimes have reason to generate that information through investigations that may turn up some potentially useful tidbit

209 See generally LAWRENCE FRIEDMAN, *AMERICAN LAW IN THE 20TH CENTURY* (2002) (noting that important legal decisions in various contexts, but particularly those implicating race relations, did not produce corresponding social conformity therewith); Barry Weingast, *The Political Foundations of Democracy and the Rule of Law*, 91 AM. POL. SCI. REV. 245 (1997) (employing game theory to find that democratic stability depends on self-enforcing equilibrium).

210 See JAMES HAMILTON, *ALL THE NEWS THAT'S FIT TO SELL* 15–17 (2004) (discussing the economics of the media and their incentives against resource-intensive forms of news generation).

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.²¹¹ Even the otherwise laudable Freedom of Information Act (FOIA)²¹² 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.²¹³

Second, whether audit bureaucracies or legislators engage in audits or something similar is an empirical question. Accordingly, this Part examines the behavior of two sets of actors capable of producing the sort of accountability-enhancing information that intrusive judicial review can. The GAO and the 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, the legislature itself can use its investigative powers directly to generate information about how executive discretion is exercised.

In fact, 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. Perhaps audits of executive discretion involving random samples of discrete applications of legal authority to specific individuals or groups are regular features of the work already carried out by the GAO and Inspectors General offices. These audit bureaucracies were, after all, created to audit the government, and their jurisdiction has expanded to include investigating the management of government programs. Their activities are sometimes directed by legislators, who (in turn) can proceed with their own audits. Do they?

211 See FRIEDMAN, *supra* note 209, at 280–90 (discussing the role of civil rights court cases in generating information that can galvanize political and journalistic responses to public officials' decisions).

212 5 U.S.C. §§ 551–557 (2000).

213 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).

A. *Audit Bureaucracies Do Not Ordinarily Perform Audits*

For the audit bureaucracies the answer is generally “no.” These agencies undertake a tremendous amount of work on bureaucratic performance, which can be considered valuable either from a social welfare perspective or from a narrower, Congress-focused view emphasizing the interests of the lawmakers that created and exert recurring influence over the audit bureaucracies. The scope of their authority is quite broad. The GAO, for example, has the power to examine “all matters” relating in some way to the disbursement of public money.²¹⁴ The Inspectors General offices in federal departments have similarly broad mandates.²¹⁵ 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).

The trajectory of the audit bureaucracies has spawned a small literature on their historical origin, legal jurisdiction, organization, and culture.²¹⁶ But relatively little is known about what the reports of these audit bureaucracies are about, what methods they use to develop their analyses, whether these reports contain recommendations that agencies actually implement, and whether any of this gets media attention. These questions are 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 a recent pilot study of the audit bureaucracies, analyzing a sample of 400 Inspectors General and GAO reports issued during the last five years.²¹⁷ The reports in the sample pertain to five major government agencies with a broad spectrum of regulatory and administrative responsibilities.²¹⁸ The analysis sought to ascertain the proportion of reports auditing execu-

214 31 U.S.C. § 712(1) (2000).

215 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.).

216 See, e.g., LIGHT, *supra* note 117; ROGER R. TRASK, DEFENDER OF THE PUBLIC INTEREST: THE GENERAL ACCOUNTING OFFICE, 1921-1966 (1996).

217 See Mariano-Florentino Cuéllar, GAO and IG Report Pilot Study: A Preliminary Analysis of the Methodologies Used by the GAO and Inspectors General to “Audit” Federal Agencies (Apr. 1, 2006) (on file with author).

218 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. *Id.* At 1-2.

tive discretion, which has been defined as involving: (1) the selection of a random sample (2) of executive branch legal determinations affecting individuals or groups and (3) identifying or defining a standard against which to assess such decisions.²¹⁹

Fewer than one in five reports in the sample appear to use any process of random selection of cases 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.²²⁰ Meanwhile, less than two percent of the reports in the sample involved audits of executive discretion.²²¹ And even among these

219 *Id.* at 1.

220 For a rare example of an Inspector General report that comes close to achieving the goal of systematically analyzing executive discretion, see OFFICE OF THE INSPECTOR GEN., U.S. DEP'T OF JUSTICE, SUPPLEMENTAL REPORT ON SEPTEMBER 11 DETAINEES' ALLEGATIONS OF ABUSE AT THE METROPOLITAN DETENTION CENTER IN BROOKLYN, NEW YORK (2003). 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. *Id.* at 1. While the report does not leverage the insight that fewer cases reviewed more thoroughly can sometimes yield greater information, its searing analysis proves 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. *Id.* at 5–7. 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.

221 Cuéllar, *supra* note 217, at 21. 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 focused 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. *Id.* at 2. Of the sample of 400 reports analyzed, about sixty-eight (or seventeen percent) used some form of random sampling. *Id.* at 21. The bulk of these did so in the context of a traditional financial audit. Only seven reports in the total sample undertook audits of executive discretion. *Id.* These covered matters such as the proportion of background checks conducted by the Transportation Security Agency on prospective airport screeners that met statutory

reports, the predominant concern was with federal grants and contracting decisions. Investigative and enforcement activities were not addressed in the small proportion of reports auditing executive discretion. Some reports occasionally chronicle problems in administrative systems like those governing aviation security. Nonetheless, audits of 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.²²² 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.²²³ Determining the proportion of the audit bureaucracies' current resources that should focus on auditing executive discretion is not the focus of this project. Instead the point is to highlight the shortcomings of traditional judicial review, to suggest that audits of executive discretion—were they to occur more regularly—could substantially address the resulting gaps, and to emphasize that the costs (measured in a number of ways) associated with such audits should be lower than the costs of meaningfully expanding traditional judicial review. Whatever the value of what the audit bureaucracies do, their tendency to focus attention on traditional financial auditing and thematic management analyses has limited their capacity to resolve problems associated with the current scope of executive discretion.²²⁴

standards and the extent of FEMA payouts for certain disaster relief programs. A separate comparison sample of fifty reports randomly selected from among those dealing with departments not in the stratified sample (covering similar dates) yielded similar results. *Id.*

222 U.S. GEN. ACCOUNTING OFFICE, GAO-04-385, REPORT TO CONGRESSIONAL COMMITTEES, AVIATION SECURITY 14-17 (2004).

223 *Id.* For a study of Inspectors General revealing that only about ten percent 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).

224 Moreover, the audit bureaucracies have incentives to distort their portfolio of activities towards a focus on procurement, contracting, grants, and financial audits and away from auditing executive discretion (particularly in the absence of congressional pressure to do the latter), relative to what is likely to be most socially desirable. The former activities are more likely to produce unambiguous financial results that can more easily support arguments for preserving and expanding agency resources. The latter may exacerbate the political animosity between the audit bureaucracies and ordinary executive agencies. See *infra* Part IV. Accordingly, changing the existing distribution of activity among the audit bureaucracies to favor audits of executive dis-

The possibility that audit bureaucracies could be encouraged to perform more audits of executive discretion without entirely neglecting other issues may provoke two skeptical reactions worth discussing. First, there is the matter of whether audits of executive discretion would be salient to the public. In fact, 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.²²⁵ Nearly 1000 articles during this period appearing in the *New York Times* mention the GAO or IG offices.²²⁶ A random sample of 200 of those news stories indicates that, while only about three percent of the stories involving the GAO appear on page one, about ten percent of those mentioning the IG offices do so. Audit bureaucracies are also discussed on broadcast news and cable channels. Even in these media (and during the same time period), nearly 500 news segments mention the various IG offices, and about thirty mention the GAO.²²⁷ These figures exceed the number of mentions garnered by many federal cabinet agencies.²²⁸

Second, the question arises whether an agency would merely ignore the prescriptive implications of an audit revealing unfavorable performance. If it 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 offices, I have already gathered data on all of the approximately 10,000 discrete recommendations made by the GAO over the last fifteen years. The data cover recommendations to the full range of federal government agencies, from the Interior Department 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 conducts follow-up interviews, additional investigations, document reviews, and issue queries

cretion is likely to yield a more defensible distribution of activity from a social welfare perspective.

²²⁵ A search ran on October 26, 2006 on Westlaw's *New York Times* database for the time period of January 2002 to January 2005 using terms "Inspector General' or 'GAO' or 'General Accounting Office' or 'Government Accountability Office'" resulted in 1354 documents. The same search ran in Westlaw's United States Broadcast Transcripts database resulted in 528 documents.

²²⁶ See *supra* note 225.

²²⁷ *Id.*

²²⁸ For example, the Department of Energy is mentioned only half as often (230 times) during the same period in broadcast transcripts; the Department of Transportation is mentioned only in 181 transcripts. Many other cabinet agencies are mentioned even less often.

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.²²⁹ Nonetheless, about seventy-nine percent of recommendations made between October 6, 1989 and February 3, 2005 were implemented, perhaps in part because of the potential media attention reports generate.

These implementation results must be interpreted with a measure of sagacity. 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 IV, I suggest some of the reasons why, as with the audit bureaucracies, there seem to be so few audits of executive discretion. In what follows I 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 tend to prioritize the investigation and monitoring of executive bureaucracies. They 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

²²⁹ The agency (quite plausibly) assumes that if a recommendation is not implemented within four years of being made, it probably will not be. Telephone Interview with U.S. Gen. Accounting Office official, in Wash., D.C. (Jan. 6, 2005).

rewards from oversight activity. Oversight allows them to 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.²³⁰ In the mid-1980s, political scientists Matthew McCubbins and Thomas Schwartz introduced what has become an incredibly durable framework for thinking about legislative oversight of the bureaucracy.²³¹ 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.²³² In contrast, “fire alarm” oversight requires less constant attention from legislators and their staff.²³³ 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).²³⁴ To encourage this sort of activity, legislators create procedures such as the federal Administrative Procedure Act²³⁵ and the Freedom of Information Act²³⁶ 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: (1) reliance on interest groups (or, on occasion, a politically engaged citizenry), and (2) episodic legislative responses to instances where these groups express profound concern with some aspect of bureaucratic activity.

Typically, legislators concerned with overseeing a bureaucracy benefit substantially from fire alarm oversight. Whatever substantive goals lawmakers may choose to pursue, the responses of organized interests constitute valuable information to all but the most politically secure legislators. But useful as it is from a narrow political perspective, fire alarm oversight is precisely the opposite of a random audit. Unless legislators directly create a procedure to audit targeted discre-

230 See JAMES Q. WILSON, *BUREAUCRACY* 235–44 (1989).

231 See Mathew McCubbins & Thomas Schwartz, *Congressional Oversight Overlooked: Police Patrols Versus Fire Alarms*, 28 AM. J. POL. SCI. 165, 166–68 (1984).

232 *Id.* at 166.

233 *Id.*

234 *Id.*

235 5 U.S.C. §§ 551–559, 701–706 (2000 & Supp. IV 2004).

236 *Id.* § 551–557.

tion (they have not 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 is the GAO and the IG offices, not their own staff.²³⁷

This is not to say that congressional investigations lack any value as vehicles for legislative oversight, or even as contributors to social welfare. 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. First, 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 are doing in the decisions more likely to be audited. Cases that are not reviewed at all do not become the subject of any legislative, political, or public pressure. Second, 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 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.

237 See, e.g., LIGHT, *supra* note 117, at 42.

IV. WHY AUDITS FACE CONSTRAINTS

Judicial review is likely to do a poor job of generating information useful in the oversight of executive discretion, because routine executive branch discretionary decisions are often committed to agency discretion or reviewed at such a low level of stringency that there is little or no meaningful chance of discovering problems. From a social welfare perspective, supplementing judicial review with audits is valuable. Audits could improve our ability to monitor bureaucracies at a reasonable cost. Even acknowledging that legislators and executive branch officials are often divided among themselves and from each other with respect to what constitutes effective performance, audits may reduce the risk of errors and bureaucratic failures that dominant political coalitions would like to prevent. Yet, audits of executive discretion rarely happen. Why?

A. *Political Actors Have Incentives to Undermine Audits*

Two important sets of actors who have stakes in the work of the executive branch may find that 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.²³⁸

On rare occasions, formal audits of executive discretion develop a political constituency within the executive branch.²³⁹ But this state of affairs rarely occurs. Instead, other things being equal, executive branch officials should react adversely to the prospect of parting with discretion.²⁴⁰ Discretion helps authorities carry out the functions that

238 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 et al., *Political Polarization and Income Inequality* 5–11 (Jan. 27, 2003) (on file with author).

239 See *supra* Part I.D.1.

240 Sometimes things are *not* equal, as when leaders in the executive branch seem to avoid responsibility for politically controversial functions holding little prospect of generating a public reward, such as choosing a site for a nuclear waste dump.

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.²⁴¹ 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 to limit their flexibility in order to demonstrate their competence,²⁴² 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.

The political viability of audits in the legislature is far from assured. It depends on the extent to which legislators with divided goals nonetheless believe that audits are worth the political price because they can uncover errors of shared concern. Such coalitions are not common, but neither are they impossible. In Congress, lawmakers who differed sharply regarding the re-authorization of the Patriot Act's investigative provisions nonetheless appeared to share an interest in limiting the most egregious law enforcement abuses of certain powers.²⁴³ Such shared concerns help explain why the most recent revision of the Act includes a rare audit provision applying to federal law enforcement information requests using the Act's subpoena and national security letter powers.²⁴⁴

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

242 For a discussion of the possibility that executive authorities could limit their discretion as a means of sending a costly signal of their competence to the public, see Cuéllar, *supra* note 139, at 57–58.

243 The legislators favoring the investigative provisions limit the risk that the authority will be used in an embarrassing manner. The legislators opposed obtain a partial victory valued by supporting constituencies and a means of generating information that could change the course of future political debates about the relevant legal provisions. See generally Bendor & Moe, *supra* note 173.

244 See Michael Sandler, *Proposed Changes to the Anti-Terrorism Law*, CONG. Q. WKLY. REP., Dec. 26, 2005, at 3394.

Audits fare less well where legislators find their political goals and the production of information through audits to be in tension. 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 is 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.²⁴⁵ 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 v. Rumsfeld* decision²⁴⁶ 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.²⁴⁷ 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.

245 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.

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

247 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* note 164.

In short, while political coalitions supporting audits are not impossible to achieve, polarization among political advocates seeking maximal advantage in their efforts to expand or limit judicial review diminishes the extent of political interest in review mechanisms that may be socially optimal. When players have more polarized views about executive branch power, substantive policy and law, or both, they probably have less to gain from investing in 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 executive discretion, and least with those for whom such a syncretic exercise seems unnecessary.

B. Institutional Inertia May Exaggerate the Value 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 do not 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.²⁴⁸

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.²⁴⁹ 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

²⁴⁸ See Daryl J. Levinson, *Rights Essentialism and Remedial Equilibration*, 99 COLUM. L. REV. 857, 926 (1999).

²⁴⁹ See *supra* Part I.C.

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 Inspectors General offices recently. Although *Vermont Yankee Nuclear Power Corp. v. Natural Resources Defense Council, Inc.*²⁵⁰ 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.²⁵¹

Indeed, while *Vermont Yankee* may arguably limit courts' abilities to directly impose audit requirements through expansive interpretations of the Administrative Procedure Act,²⁵² audits seem directly germane to the familiar procedural due process balancing framework ordinarily traced to *Mathews v. Eldridge*.²⁵³ By privileging the importance of accuracy-enhancing mechanisms, *Eldridge* 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.²⁵⁴ As they respond to conceptual and political objections, auditors (and audit supporters) may thus find solid ground from which to empha-

250 435 U.S. 519 (1978).

251 Cf. Bressman, *supra* note 2, at 555 (laying out a new theory for "examining more directly the concern for arbitrariness").

252 *Vermont Yankee*, 435 U.S. at 524 (holding that courts are generally not free to impose additional procedural requirements beyond those established by the APA).

253 424 U.S. 319, 323–28 (1976).

254 See *id.* at 334–35 (establishing a three-part test focusing on increasing accuracy in administrative decisionmaking). Indeed, the frequent relative absence of an empirical basis informing courts' determinations of what process is due in *Eldridge*-like cases raises the question of how courts are currently making these determinations, and how these might be best evaluated.

size the potential intersections between the work of courts and auditors.

C. *Audit Bureaucracies Define Their Mandates Narrowly*

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."²⁵⁵

Politics has cemented the early focus on financial auditing and procurement.²⁵⁶ 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 improprieties that, where rectified, produce additional government

²⁵⁵ LIGHT, *supra* note 117, at 42 (quoting GEN. ACCOUNTING OFFICE, FINANCIAL AUDITS IN FEDERAL EXECUTIVE BRANCH AGENCIES (1978)).

²⁵⁶ *Id.* at 43-45.

revenues.²⁵⁷ 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 proves difficult to find legislators or executive branch officials in favor of waste, fraud, or (financial) abuse—though (particularly for the GAO) it is 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.²⁵⁸

On occasion, legislators may find audit bureaucracies useful to generate publicity and promote policy objectives not related to advancing the ubiquitous mantra of eliminating waste, fraud, and abuse.²⁵⁹ Nonetheless, the legislative requests that drive a considerable proportion of the GAO's work (and probably some of the work of Inspectors 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.²⁶⁰

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

257 See, e.g., *Seven Years of GPRA: Has the Results Act Provided Results?* 106th Cong. (2000) (testimony of the Hon. Pete Sessions), available at 2000 WL 1211285 ("GAO has identified waste, fraud and abuse in government programs in amount of over \$800 Billion in the last 5 years.").

258 See, e.g., Anne Margaret Joseph, *Political Appointees and Auditors of Politics* 8 (May 2002) (unpublished Ph.D. thesis, Harvard University) (on file with author).

259 See LIGHT, *supra* note 117, at 44–45 (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.").

260 *Id.* at 84–85 (providing a table summarizing the backgrounds of IGs during the Carter and Reagan administrations).

certainly leave the audit bureaucracies with a measure of discretion.²⁶¹ 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)²⁶² 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 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.²⁶³ 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

261 See, e.g., Joseph, *supra* note 258, at 146 ("Auditors—whether in economic or political institutions—choose what to observe and report in order to advance their objectives.").

262 *Id.* at 179 (statutory duties); *id.* 181–82 (describing the areas of work of the GAO); *id.* 209 (describing the high portion of reports requested by Congress in recent decades).

263 Cf. DANIEL CARPENTER, *THE FORGING OF BUREAUCRATIC AUTONOMY* 353 (2001) (noting that "[b]ureaucratic autonomy is politically forged"). 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). *Id.* at 353–54. 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. Cf. Mariano-Florentino Cuéllar, *Refugee Security and the Organizational Logic of Legal Mandates*, 38 *Geo. J. INT'L L.* (forthcoming Nov. 2006).

the larger public) that its autonomy should be protected.²⁶⁴ The development of public bureaucracies during the last century provides ample evidence not only 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.²⁶⁵ At the Postal Service,²⁶⁶ the Forest Service,²⁶⁷ and the Food and Drug Administration,²⁶⁸ 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.²⁶⁹ 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

264 The example of independent electoral and redistricting commissions in other countries is instructive. As Pildes notes, a number of democracies, including the United Kingdom, Australia, Canada, and New Zealand, use nonpartisan commissions to draw political boundaries, where “politicians typically have no role.” See Richard Pildes, *The Constitutionalization of Democratic Politics*, 118 HARV. L. REV. 28, 78 n.211 (2004). Given that these commissions exist in the context of a democratic political system, political actors may face electoral sanctions offsetting the advantages of interfering with these commissions. See also Christopher Elmendorf, *Representation Reinforcement Through Advisory Commissions: The Case of Election Law*, 80 N.Y.U. L. REV. 1366, 1432 (2005) (“The AC [Electoral Advisory Commission] that wants to enact a reform that most legislators disfavor must establish a base of public support for its plan.”). 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.

265 See CARPENTER, *supra* note 263 (providing a series of specific examples of agency involvement in policymaking).

266 *Id.* at 123 (discussing the implementation of rural free delivery).

267 *Id.* at 282 (discussing vast changes within the Forest Service policy without congressional approval).

268 *Id.* at 366 (discussing the Food and Drug Administration’s “tremendous power over the pharmaceutical industry”).

269 See also Daniel Carpenter & Colin D. Moore, *Robust Action and the Strategic Ambiguity in a Bureaucratic Cohort: FDA Officers and the Evolution of New Drug Regulations, 1950–1970*, at 8 (Harvard Univ. Dep’t of Gov’t, Working Paper, 2006) (“If rulemaking officials (or cohorts) are able to secure public legitimacy, are able to assemble “strange bedfellows” coalitions behind their activities, and are able to capitalize upon political opportunities, they can achieve the translation of divergent bureaucratic goals into policy while bypassing entirely the legislative process.”).

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 is what lets government protect the environment, keep workers safe at industrial sites, and fight military 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,²⁷⁰ or how Hoover's FBI libelously fed speculation that slain civil rights workers were promiscuous and mentally ill subversives,²⁷¹ 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 three implications of these facts. First, 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 face greater difficulties learning without external mechanisms to systematically review and critique their work. If additional resources to monitor high-discretion activities were available, it would be better from a social welfare perspective that we audit instead of dividing them across an entire pool of cases. Indeed, the mere presence of judicial review can foment the misleading impression that external review is succeeding in restraining abuse and excess, when in fact the probability of discovering problems may be close to zero.

270 See James L. Sundquist, *Reflections on Watergate: Lessons for Public Administration*, 34 PUB. ADMIN. REV. 453, 456 (1974).

271 See REACHING FOR GLORY 245 (Michael R. Beschloss ed. 2001).

Second, the conventional focus on traditional judicial review as the preeminent means of supervising executive branch discretion foments another deceptive impression: that problems with imposing more stringent judicial review mean executive discretion should be relatively immune from formal review. 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 involves *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. Though not invariably, this state of affairs often proves acceptable for presidential administrations, executive officials, and legislators supporting executive power: they will tend to be perfectly satisfied allowing that power to evade more frequent review. Somewhat counterintuitively, 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.

Third, a discussion of auditing serves a more immediate function—by demonstrating how arguments for broad discretion are often radically underdeveloped. Cogent arguments in favor of limiting interference with executive discretion should not herald the end of a discussion, but its beginning. There remains the task of balancing benefit against costs, and deciding on the institutional rules that should govern the relationship between courts, agencies, audit bureaucracies such as the GAO and Inspectors General, legislatures, and external interests. As one viable approach to structuring those institutional rules, audits help elucidate the conceptual murkiness of common arguments that favor broad executive discretion without considering how to enhance mechanisms producing information about executive branch performance. If executive discretion is to be defended coherently, the case for it ought to transcend reflexive critiques of judicial supervision, or veneration of a political process whose very impact is contingent on institutional mechanisms to evaluate bureaucratic performance.

It should be self-evident that the existence of broad discretion is not synonymous with poor bureaucratic performance. When coupled with an appropriate mix of external political constraints and intrinsic personal motivations, such discretion can produce a harvest of valuable activity from regulatory, criminal justice, and national security bureaucracies. What borders on madness is to assume that executive decisions of such profound importance—such as whose assets to freeze, whom to indict, or against whom to enforce environmental regulations—will routinely advance some defensible conception of societal interests when existing law lets them so readily elude scrutiny. In effect, the challenge of calibrating the law’s relationship to bureaucratic discretion constantly reiterates a set of complicated institutional design questions about the use of public power. Arguments for executive power that ignore the subtleties of institutional design become

nothing more than Trojan horses, with the contraband assumptions hidden inside meriting the same careful scrutiny and critical judgment that executive bureaucracies themselves deserve.

