CONTESTED CONSTRAINTS:
REGULATORY STATUTES IN AMERICA’S
MODERN ADMINISTRATIVE STATE

Philip Alexander Wallach

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF
POLITICS
Adviser: Keith E. Whittington

September 2012
Dissertation Abstract

Leading models of bureaucratic delegation assume that congressional statutes create clear and effective boundaries that constrain executive branch discretion under the law. Focusing on regulatory statutes, this dissertation investigates the nature of these constraints, the mechanisms through which they are realized and made effective in practice, and various conditions that can lead to their failure. In doing so, it highlights the central role of contesting legal interpretations in determining statutes’ functional meaning, the irreducibly verbal nature of statutory interpretation, and the difficulty of the task facing interpreters.

Drawing on analogies to constitutional and contractual interpretation, the dissertation explains what statutory text can and cannot do. Truly clear text commands near-universal obedience from interpreters in both the executive and judicial branches—and in these instances, principal-agent models depicting Congress as conferring instructions and bureaucrats and judges as carrying them out are largely adequate. But legislators often draft statutes that centrally feature indeterminate language. Where text is indeterminate those charged with making sense of it define its meaning through usage and contestation. These acts of construction are rooted in, but not determined by, the content of the text and necessarily incorporate values not found within the text. Conceptualizing the statute’s influence in terms of a principal-agent dynamic is therefore problematic.

Through in-depth analysis of several original case studies (the history of the Glass-Steagall Act’s interpretation; the Food and Drug Administration’s attempt to regulate tobacco during the 1990s; the application of the Clean Air Act to greenhouse gasses) and an original large-n dataset (rulemaking under the Clean Air Act in the 1990s and 2000s), the dissertation
traces the processes through which textual constraints become binding. A major theme that emerges is that judges take into account the policy-specific context in which novel interpretations occur, and especially the relative institutional capabilities of Congress and bureaucrats. Where Congress seems incapable of providing guidance, judges are more likely to accept strained interpretations of statutory text.

The dissertation’s most important contribution is to raise neglected empirical questions about the effects of law, answers to which would be of direct use to policymakers.
Acknowledgements

Although contemporary political science has given me many opportunities to gripe—and I have passed up few—it would pain me to leave the impression of being a malcontent. My time as a graduate student at Princeton has been deeply enriching, satisfying, and happy, for which I owe many people thanks.

First and foremost are my advisors, who have been at the heart of my experience. Neither Keith Whittington nor Doug Arnold is easy to impress, and I have benefited enormously from their unsentimental criticism. Keith’s work ethic and remarkable sense of objectivity have set an example I aspire to follow, and his ideas have left a deep imprint on my thinking. Doug’s patience, reliability, and willingness to share his long experience have been invaluable assets, and nobody could ask for a more careful or thoughtful editor.

Many other professors have also helped me along the way. Dave Lewis graciously found time for me after moving to Vanderbilt, and his comments have greatly strengthened this work. Chris Achen provided me support throughout my time at Princeton and helped me to develop strong common-sense ideas about data analysis. Julian Zelizer was a valuable mentor and a slightly intimidating example of what a publicly-engaged academic career can be. Paul Frymer and John Kastellec, whose courses I had the pleasure of precepting, both provided excellent feedback as my project developed. Chuck Cameron was especially helpful in forcing me to think through my theoretical perspective. Kim Scheppele gave helpful feedback, and her leadership of the Law and Public Affairs program improved my Princeton experience in countless ways. Robbie George contributed to my thinking through his excellent, Oxford-centric Philosophy of Law seminar and his leadership of the James Madison Program in American Ideals and Institutions, which consistently brought stimulating speakers to campus. Two previous
intellectual rabbis, Richie Adelstein of Wesleyan and Michael Greve of AEI, have continued to offer their wisdom and encouragement as I’ve progressed.

I’m grateful to Princeton’s law librarian, David Hollander, who was always ready to help figure out primary source materials.

My classmates at Princeton have also been instrumental to my work. Emily Zackin provided invaluable counsel in helping me to figure out what I wanted to do. Herschel Nachlis and Sean Beienberg each read multiple chapters of the manuscript and provided insights far more valuable than they are willing to admit. Herschel has also served as confidante, confessor, and friend during my years at Princeton, and has done much to keep me sane. I’ve also benefited greatly from feedback from Alex Acs, Deborah Beim, Nick Carnes, and Alex Ruder. I’m also grateful for all of the comments I received through the American Politics Research Seminar and the Public Law Working Group.

Interfacing with the legal world at various points in my project also proved helpful. Discussions with Mike Livermore, Peter Strauss, and especially Judge Bob Katzmann helped me to clarify my thinking and provided encouragement. Several lawyer friends also provided much-needed outside perspective on Chapter Four: Justin Gundlach, Zeke Hill, and Rob Weinstock. Lynda Dodd also provided thoughtful feedback on the chapter as a commentator at the Midwest Political Science Association meeting.

Former U.S. Representative Jim Leach graciously allowed me to interview him about the long downfall of the Glass-Steagall Act, and Chapter Five is much stronger for his contribution.

I should recognize two institutions that have provided material and intellectual support for my work. I have been honored to be chosen by the Institute for Humane Studies as a Humane Studies Fellow throughout my graduate career, and their programs have consistently
contributed to my professional development and the evolution of my ideas. Program Director Nigel Ashford has been an invaluable guide to navigating graduate school. In my final year at Princeton, I was also fortunate to be a part of the Fellowship of Woodrow Wilson Scholars, which provided a much-appreciated opportunity to interact with graduate students from other disciplines.

Finally, I thank my family. My brother-in-law, Serge Krimnus, has challenged my thinking in delightfully unexpected ways, and I look forward to one day winning an argument with him. My sister, Kerry Wallach, led by example in becoming Dr. Wallach, and her devotion to her work is an inspiration. My mother, father, step-mother, grandparents, mother-in-law, and grandmother-in-law all quietly expect and believe that I can accomplish important things through my work, and I draw strength from their confidence, love, and support. Most of all I thank my wife, Vera Krimnus. She contributed to this project directly and indirectly in countless ways, the most important being that she makes me smile each and every day.
Table of Contents

Abstract iii

Acknowledgements v

Chapter One – 1
Introduction

Chapter Two – 54
Analogizing Statutes, Contracts, and Constitutions

Chapter Three – 116
The Grammar of Legal Determinacy, Contestation, and Constraint

Chapter Four – 179
When Can You Teach an Old Law New Tricks?

Chapter Five – 251
Statutory Erosion and Maintenance:
The Case of the Glass-Steagall Act
Chapter One – Introduction

“[Always] there is this restless questing: what difference does statute, or rule, or court-decision, make? […] This is a question in first instance of fact: what does law do, to people, or for people? In the second instance, it is a question of ends: what ought law to do, to people, or for them? But there is no reaching a judgment as to whether any specific part of present law does what it ought, until you can first answer what it is doing now. To see this, and to be ignorant of the answer, is to start fermenting, is to start trying to find out”
—Karl Llewellyn, “Some Realism About Realism” (1931, 1222-23)

An “orgy of statute making” (Gilmore 1977, 95) has transformed the American legal landscape over the past century, so that today America’s economy is blanketed by statutory law at both the state and federal levels. This dissertation asks how the words of federal regulatory statutes are transformed into meaningful constraints on government actors in America’s administrative state through processes of interpretation and contestation.¹ In other words, I examine laws as determinants of policies rather than as outputs of the political process.

Adopting this perspective leads me to focus on law’s verbal nature as well as the complexity of legal interpretation. This often stands in contrast to scholars of law-as-output, who can afford to treat the concepts of “law” and “effective legal constraints on policy” as interchangeable. In addressing their research questions, they have many good reasons to make such simplifying assumptions, and by viewing statutes as instruments of control for a legislative principal seeking to control bureaucratic agents they have produced many important insights.² But it is important to question when their assumptions will hold and, further, how they can fail.

¹ While I focus on American administrative law, many of my arguments and conclusions should apply with equal force to other areas of law and policy where statutory text plays a central role. Similarly, the logic employed in this dissertation strikes me as potentially applicable to regulatory law in other nations, although different sets of institutional structures could support very different orientations toward statutory text; see Sunstein (1999, 658-661) for speculation along these lines, suggesting that British formalism might be stronger than American because its laws are more easily revised by Parliament.
² This is especially so when they think about how the broad outlines of legislation are determined by institutional designers. It is no coincidence that many of the leading contemporary separation-of-powers models focus on the moment of agency genesis, e.g., Epstein and O’Halloran (1998), Lewis (2003), or the passage of new statutes, e.g., Huber and Shipan (2002).
Examining law in action for these purposes leads me to stress that carrying out ambiguous statutory commands is necessarily difficult and contentious. Officials charged with carrying out the law must attempt to conform to the dictates of statutory text while achieving a fair result, preserving the coherence and stability of the legal edifice, and supporting a policy scheme that is workable in practice and effective in realizing worthy goals. It is this precarious balancing act that determines whether our laws and policymaking system will “work,” in the sense of avoiding malfunctions both needling and catastrophic.

As the opening quotation from Llewellyn (1931) makes clear, understanding how law works is a prerequisite for being able to say how to make it work better. As I discuss in greater length toward the end of this Introduction, one of this dissertation’s primary objectives is to illuminate a research agenda that speaks directly to policymakers, aiding them as they make particular choices in legislative drafting or in statutory interpretation in ways that more abstract inquiries are incapable of doing. I hope to convince political scientists, whose sense of objectivity and methodological sophistication are in short supply in studying the workings of the law, that this research agenda represents an important opportunity to make the law better.

In this Introduction I proceed to consider existing models of thinking about the role of law in America’s administrative state. As shown in Table 1, I sort these models into four cells, based on two dimensions. First, does the model treat law as simple and self-interpreting, or as complex and legitimately contentious? Second, does the model conceive of law entirely in terms of a principal-agent relationship?

As I proceed, I explain why those models treating law as simple as well as those thinking entirely in principal-agent terms are inadequate for my purpose of thinking realistically about

---

3 Although I discuss model-types, I make no claim to an exhaustive cataloging here.
law’s ability to constrain or guide governmental action. In doing so, I necessarily adopt a critical tone. One might fairly argue, though, that the models I discuss do not deserve to be criticized for failing to accomplish a goal that they never intended to accomplish. After all, models are all incomplete and are best conceived of as purpose-sensitive (Clarke and Primo 2011), and so it makes little sense to criticize them for failing to capture everything.

This point is well-taken—at least in theory. Users of models should make their models work for them, rather than the other way around, and if they all did so then there would be little need to harp on the point that each model is limited to a particular sphere. Modelers often make far more ambitious claims on behalf of their creations, though. Rather than offering their models as helpful in thinking about some conceptual problems, they insist that their model presents a superior way of thinking about reality. As an afterthought, they admit that there are other factors at work, but they largely dismiss these as marginal phenomena. This description is especially true of many game theoretic modelers in the rational choice tradition, especially those who attempt to show that their models effectively predict observed outcomes (sometimes labeled EITM, or Empirical Implications of Theoretical Models). Rather than placing the limits of their modeling efforts front and center, many scholars are instead at pains to put their models in a “horse race” with “conventional” ways of thinking, thereby showing the superior insight of their own contribution.

---

4 See Moe (2009) for a rousing celebration of this school’s achievements in the field of Presidency studies, and Skowronek (2009) for an effective, deflationary rejoinder, emphasizing the limits of inquiry that the rational choice framework seems to impose on its users without their full awareness. See also Clarke and Primo (2011) for a deeper critique of the EITM program focusing on its logical difficulties and its grounding in old-fashioned philosophy of science.

5 Along these lines, many judicial politics articles contrast the explanatory power of the attitudinal, strategic, and “legal” models. As Friedman and Martin (2011) point out, devising some tests in which the legal model (which often doesn’t look like a coherent model at all) performs poorly and then inferring that law’s overall importance is minimal is a classic non-sequitur.
Although the term “model” is less frequently used, those in the legal academic world can also become quite convinced that their theoretical frameworks provide comprehensive and complete accounts of reality—or at least that they should. With important exceptions, textualists often talk as if their creed offers definite answers to every interpretive question; intentionalists and purposivists sound just as sweepingly confident in rebuttal. Acting as practitioners (rather than casual theorists) bureaucrats, lawyers, and judges are far less likely to be entirely doctrinaire (Posner 2008; Rakoff 2010), perhaps because real-world consequences follow if they over-privilege theory.

By discussing what existing models can and cannot explain, I show where this dissertation’s contributions fit into the existing literature as well as sharpening the lines of disagreement between myself and others.
<table>
<thead>
<tr>
<th>Law is</th>
<th>Limited to Principal-Agent</th>
<th>Going beyond Principal-Agent</th>
</tr>
</thead>
<tbody>
<tr>
<td>simple</td>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td></td>
<td>Mechanical law – I.A</td>
<td>Clear legal limits, then policymaking – II.A</td>
</tr>
<tr>
<td></td>
<td>Vulgar formalism – I.B</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mechanical law, ignored – I.C</td>
<td>Law not pictured – II.B</td>
</tr>
<tr>
<td></td>
<td>Hybrid – I.D</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>Congressional dominance – III.A</td>
<td>Principal-agent generally misleading:</td>
</tr>
<tr>
<td></td>
<td>Mainstream legal academic debates between textualism,</td>
<td>Constraints present, but not legislative:</td>
</tr>
<tr>
<td></td>
<td>intentionalism, purposivism, etc. – III.B</td>
<td>Dworkinian judging, natural law, equity of</td>
</tr>
<tr>
<td></td>
<td></td>
<td>the statute – IV.A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Skeptical legal realism – IV.B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No real constraints:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Radical indeterminacy – IV.C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Principal-agent mostly right, but must also</td>
</tr>
<tr>
<td></td>
<td></td>
<td>go beyond:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Textualism + defer to executive – IV.D</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weak textualism + defer to legislature or</td>
</tr>
<tr>
<td></td>
<td></td>
<td>judiciary – IV.E</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weak textualism + muddle through – IV.F</td>
</tr>
</tbody>
</table>
Cell I – Simple law, wholly within principal-agent framework

If we assume that a) the principal (legislature) has exhaustive preferences about all aspects of policy questions, and b) it is capable of perfectly articulating its views without serious ambiguities, then we can entertain one of the following simple views about the operation of law as a constraint.

I.A. Mechanical law: the executive branch acts as a mere “transmission belt” of the legislature’s statutory commands, with laws providing self-evident answers to every question that arises. Judges merely enforce obedience to these clear answers. In other words, laws are basically self-interpreting and self-enforcing. Courts are assumed to be ready, willing, and able to detect deviations from the law and effectively sanction them.

Many of the classic games of statutory interpretation incorporate some version of the mechanical conception into their models (e.g., Landes and Posner 1975; Ferejohn and Shipan 1990; Eskridge and Ferejohn 1992; Ferejohn and Weingast 1992a, 1992b); and it is not so uncommon for political scientists of varying methodological stripes to adopt this outlook. Some

---

6 Stewart (1975) describes a “traditional” model of administrative law in which agencies are thought of as “mere transmission belts” for legislative directives and administrative actions must take whatever form facilitates judicial oversight, so that the judiciary can easily enforce compliance with the legislature’s commands. He then goes on to describe why this model has been largely discredited and how judges have struggled to transform traditional doctrines to ensure the representativeness of administrative processes.

7 In terms of a formal mapping of real situations into legal categories, this simple view assumes that laws deal only in “crisp” mathematical sets, in which all elements are well-defined. As Lerner and Wanat (1983, 501) point out in the context of “bureaucratic categorizations,” however, policymakers often deal with “fuzzy sets,” defined as “collection[s] of elements whose membership is equivocal.”

8 To be fair, nearly all of these articles acknowledges the limitations of the mechanical view, and indeed Ferejohn and Weingast (1992, 268) recognize it as “naïve textualism.” Many of the models presented in these papers might in fact be better classified in group I.B or I.D.

9 Many authors seem to subscribe to the intellectually spurious belief that laws can effectively bind bureaucrats to a single outcome simply because they are laws. Llewellyn (1930, 449) could already declare, some eighty years ago, that “the ancient assumption that law is because law is” was in eclipse, and yet adherence to this view has remained remarkably persistent. One clear example comes from the otherwise illuminating work of Kirst (1969, 64), who expresses the not-all-that-remarkable view that “statutory controls [on bureaucracy] are self-enforcing as the law of the land.” This sentiment confuses the pervasive normative conviction that public officials should follow the law with a positive belief that the law is automatically controlling.
version of mechanical law may be the “legal model” sometimes derided by authors in the judicial politics tradition.\textsuperscript{10}

\textbf{I.B. Vulgar formalism:} Quite apart from the academic modelers, there are many people who believe that “law isn’t really that hard.” In this view, “the law” is out there, ready to provide an answer to nearly every legal question if the proper rules of legal reasoning are obeyed. There are rules, including those prescribed by statutes, and bureaucrats, judges, and citizens should be able to follow them unless they are knaves or fools. Laws do not interpret themselves—but honest interpreters \textit{ought to} reach the same understanding of statutory requirements. This position is what Leiter (2010) calls “vulgar formalism” in recognition of its wide currency among laypeople and even many lawyers and judges. Arguably, this “just follow the law” mindset is a valuable component of a functioning rule of law, and “faith” in its adequacy deserves to be taken seriously (Smith 1999). In any case, few scholars find this position to be tenable for any in-depth purposes.

Those who subscribe to the strongest forms of vulgar formalism should be most satisfied with treating the simple models as if they get things mostly correct. If the “rules of the game” are basically transparent, modeling law as a game of rule-following or willful rule-disobeying is appropriate. Vulgar formalists basically argue that judging \textit{can and should} be more like the models described in I.A, such that “the perplexities of statutory interpretation [are] due merely to legal sophistry” (Farber 1992, 534). Because this view tends to dismiss examples of complexity as impious subversions, someone stubbornly defending this position in a lawyerly way is likely to create a well-fortified, if not terribly enlightening, theoretical edifice.

\textsuperscript{10} Friedman and Martin (2011) argue that this “legal model” as presented is often no model at all, but rather just a disorganized cluster of ideas about what judges are said to care about.
I.C. Simple cynical view – ignoring law: mechanical application of the laws is possible, and indeed just as easy as I.A suggests, but policy actors (including judges) generally ignore or circumvent any legal constraints, making decisions based on their own ideological preferences. When scholars bother to think about the executive branch’s relationship to law (as opposed to ignoring it – see II.B), few think that executive branch actors merely disregard clear legal directives; for all the talk of an imperial Presidency and “delegation run riot,” most scholars of bureaucratic action believe that law acts as some sort of constraint on or basis for executive branch decision-making. In contrast, there are a number of scholars who have argued that judges, and especially Supreme Court Justices, are almost entirely unbound by law, leaving them to act as “Legislators in Robes” (Segal and Westerland 2011). Although such broad statements of judicial limitlessness tend to be inspired by cases of constitutional judicial review rather than statutory interpretation, claims are often made that include statutory interpretation cases. Judges as maximizers of exogenous preferences appear in countless game theoretic models in the judicial politics literature, even when the subject is statutory interpretation.

11 The phrase is from Lowi (1979), who was concerned with the extent to which Congress had willingly ceded policymaking authority to the executive branch through its enactments (a complaint that presupposes that laws could constrain bureaucrats effectively if only they tried to do so).

12 Perhaps the closest to being an exception are Posner and Vermeule (2010), who argue that “liberal legalism” and its faith in effective legal constraints on the executive branch is nearly delusional, given the extent of the branch’s advantages in terms of expertise and decisiveness. Often, though, their discussion fails to match the strongest articulations of their thesis. For instance, when an executive branch official searches for a power to justify some action, the authors claim that “there is usually some statute or other in the picture to which the president or an agency can plausibly appeal” (11). This seems true enough—but “usually” and “always” are miles apart given their claims of inevitability. They dismiss situations in which law seems to provide a binding constraint on executive choices as mere “palliatives” (28), but it seems that the issue is precisely whether, when, and how such limits might be effective (even if, as they argue, in some cases efficacy might be impossible).

13 Admittedly, thinking about Supreme Court Justices as straightforward political actors makes some sense if the field of inquiry is restricted to the most politicized and controversial cases; hence the near-unvarying focus on the Supreme Court’s civil liberties docket by the authors who insist on this view. Even if one turns to the Supreme Court’s whole case load—which of course only represents a tiny fraction of the overall business of the federal courts—this picture begins to look implausible in light of the prevalence of unanimous or near-unanimous opinions.

14 One might argue that I.B belongs on the right side of Table 1, since in these models judges act to maximize the satisfaction of their own preferences, rather than working as agents of the legislature. But this is characteristic of every principal-agent problem: the whole point is that unless the principal has in place some incentive-compatible
I.D. Simple semi-cynical: positions I.A and I.C could be combined, such that judges and bureaucrats make decisions based in part based on their own (exogenous) preferences and in part on the requirements transparently set by law. This is a natural interpretation of the way law is modeled in Huber and Shipan (2002), where an agency faces a probabilistic chance of judicial reproach if they stray outside the limits set by the law. This stochastic process might represent uncertainty as to whether judges would act as faithful enforcers of the law or as ideological sympathizers with the agency’s extralegal action.\textsuperscript{15}

What are the limitations of these “simple” models?

Each of these models makes the strong assumption that every legal interpretation can be easily identified as consistent or inconsistent with the law—by bureaucrats, judges, and regulated interests alike (or at least these parties’ lawyers). For game theoretic modelers, this simplifying move is made through the rendering of law and policy in spatial terms (most often in a single dimension); every interpretation (a term which becomes synonymous with “policy”) is either within the law’s set of permissible actions or it isn’t.

\textsuperscript{15} Alternatively, this modeling choice could be taken to represent a more subtle conception of the law, where the supposed legal limits capture only what seems to be impermissible, thus preserving uncertainty about the law’s meaning. This interpretation fits uneasily with Huber and Shipan’s (2002) modeling choices, though. The degree of uncertainty (represented by $\gamma$), is modeled as exogenous and constant regardless of the action the agency chooses, rather than endogenous to the type or magnitude of non-compliance. An action within the compliance window has no chance of being reversed; and two actions, one just outside of that window and one far beyond it, have an equal chance of reversal. In this modeling, there can be no sense of legal outlandishness; no boundary-transgressing is any more controversial than any other.

A modeling option that would better capture the idea of legal ambiguity would be to have the legislature choose a point of “core meaning” for the statute, and make bureaucratically-chosen policies less likely to withstand legal scrutiny the farther they are from that point—and then dispense with the idea of a window altogether. Mathews (forthcoming 2013) provides a model in this spirit, although not framing things in exactly these terms. In general, this seems like a promising modeling technique that should be explored more.
In all but a few unusual cases, this geometric clarity is a poor approximation of reality. Instead of a clear spatial meaning of the underlying statute, all anyone has in front of them is a bunch of words. Words can be both less restrictive (because of the possibilities for creative interpretation) and more restrictive (because of the many essentially non-ideological practical requirements that following statutory language entails) than a simple delegation of a clearly bounded one-dimensional choice would be. Words create idiosyncratic challenges that lawyers and judges take seriously and that do not map cleanly onto ideological space. Administrative law opinions frequently hinge on the definition of a key word as used in several different paragraphs of the law; or whether a phrase like “comparable to” has a determinate meaning in a particular context or is open to agency interpretation. For those affected by the laws and for the policymakers who administer them, many of these questions are irreducibly verbal, and have answers that must be arrived at through interrogating the possible meanings of words. In many cases, it would be impossible to determine which arguments belonged to which side if divorced from the context. The regularity of unanimity on Circuit Court panels and Supreme Court statutory interpretation decisions strongly suggests that it is the exception in which the play of words becomes a mere smokescreen for an ideological power struggle. (Which is not to say that this never happens—surely it does.)

When thinking about the effects of statutory language, this dissertation tries to avoid both simplistic extremes; legal language is neither self-interpreting and automatically controlling nor an irrelevancy disregarded by fundamentally political actors. Chapter Two explains the content of this “middle way” at length, but it can be readily summarized: clear law can be expected to

---

16 The exception would be a law explicitly requiring choices of a continuous (probably numeric) variable, such as a law of the form: “the Secretary shall set a limit of between X and Y parts per million of pollutant in the atmosphere, as measured with technique Z.” This is indeed a facially clear law, in need of very little interpretation; setting a limit less than X or more than Y is clearly impermissible. Few laws do follow this form, however.
control government actors’ choices, but determining what statutory language is really clear is
difficult; ambiguous, indeterminate law can be expected to shape and guide government actors’
choices, but it cannot be expected to decisively resolve every subsidiary question. Because real
statutes are usually ambiguous, “following” their words can often be very difficult (or even
impossible), without anyone being duplicitous or incompetent. Bureaucrats, judges, and those
regulated by the law take statutory language seriously—but this describes the nature of their
enterprise, rather than definitively saying what they will do.

As a result, I think of the legislative process as bestowing large bundles of legal-
responsibility-creating words upon agencies, complete with (anticipated and unanticipated)
problems of interpretation, rather than as conveying well-thought-out, crystal-clear “core
meanings” or “compliance windows.” Chapter Three thinks about why legal disputes about
rulemaking arise through this verbal lens: as the agency tries to expound the meaning of its
statutory inheritance, it will tend to run into legal disagreement and contestation whenever the
language was clear enough to imply that some constraint on their actions was intended, but the
words used to frame that constraint were not self-evident. The practical meaning of such
requirements will be defined through practice and contestation.

**Cell II – Simple law, beyond principal-agent**

Because they also reduce legal interpretation to something transparent and simple,
Models in Cell II are subject to many of the same shortcomings discussed in relation to the
simple models of Cell I. But these models do allow for complexity of another sort, by
recognizing that governmental actions are not always best explained with a principal-agent
framework.
II.A. **Clear legal limits, then policymaking**: law may set clear limits, and a principal-agent framework appropriately captures these. But legal guidance from the legislative principal runs out, and beyond its reach is a realm of policymaking unaffected by law. Arguably, one could look at some of the spatial models described in I.A or I.C in this way: law’s limits are clear and controlling, but they leave a realm of action within which the interpreter can freely maneuver. In this light, bureaucrats operating within a compliance interval are nevertheless creative forces guided by purposes independent from the legislature’s.\(^{17}\)

Although they would undoubtedly reject the conception of law as transparent presented by the models in Cell I, leading administrative law experts Bamberger and Strauss (2009, 611-14) argue that while courts should use “traditional tools” of statutory interpretation to determine if some action is clearly impermissible, “judges must withdraw to a supervisory role when agency choices fall within a zone of ambiguity left by congressional instructions.” In *Chevron*\(^{18}\) Step One, judges determine the boundaries of the statute’s ambiguity and ensure that the action before them falls within that permissible range; in *Chevron* Step Two, judges ensure that the decision-making process used to reach the construction was reasonable (not arbitrary or capricious).

I argue that it is impossible to clearly demarcate “permissible interpretation” and “reasonable policy,” with the law controlling in the former case but not in the latter. Following Stephenson and Vermeule (2009), I assert that these distinctions collapse upon close

\(^{17}\) Whether such games are thought to include some component that is “beyond” the principal-agent framework is mostly a semantic question. The judgment would probably depend on how consequential the band of discretion left open to bureaucrats is; if narrow, then it would seem rather grandiose to see the bureaucrats as something other than the legislature’s agent; if broad, then it is customary to speak of a government agency wielding “delegated” powers as doing something consequential and interesting in its own right. Although both the political science literature and courts considering the non-delegation doctrine often act as if there is a bright line between those laws that delegate and those that merely charge the executive branch with some responsibility, I admit that I have difficulty understanding how a clean distinction could ever be drawn.

\(^{18}\) This dissertation discusses the doctrine of *Chevron v. NRDC* (1984) many times. Chapter Two, Sections 3.A and 4.A discuss the doctrine’s logic in detail.
examination, such that the primary question facing courts is always whether an agency’s actions can be seen as consistent with its statutory mandate.\textsuperscript{19} \textsuperscript{20} As a result, in the discussion of Cell IV below, I emphasize that law continues to guide actions under a statute even when it does not require any particular outcome. This claim is more or less nonsensical if we conceptualize legal requirements in low-dimensional spatial terms, but has important implications when thinking of law as made up of verbally formulated requirements.

II.B. Law not pictured: Of course, in many models of government action, law is nowhere to be found. This includes models in which judges are present as preference-seeking actors without any reference to what the law might require of them. Such models offer little to analyze in terms of law’s effects. I want to clearly note, however, that someone committed to game theoretic modeling of government action could fairly point out that their methodological approach can capture both principal-agent dynamics of legal control and extra-legal strategic actions that are outside of a principal-agent framework, albeit not in the same model.

\textsuperscript{19} At the same time, their suggestion that \textit{Chevron} be treated as having only a single step is unhelpful; it is still useful to first address the question of whether the statute clearly prohibits am action (or inaction) before moving on to ask whether an agency has acted in a way that is broadly consistent with statutory requirements that are admittedly indeterminate.

\textsuperscript{20} The Supreme Court’s language in \textit{Chevron} supports this position. Step Two does not ask about reasonableness in the abstract, but “whether the agency’s answer is based on a permissible construction of the statute” (467 U.S. 837, 843). Gap-filling regulations are to be prohibited if they are “manifestly contrary to the statute,” a requirement that seems to be quite redundant with Step One; similarly, where there is implicit delegation, an agency’s “reasonable interpretation” of the statute is to stand (and not just \textit{any} reasonable action) (844). This language suggests that the statute does not recede from the picture even as it fails to provide a determinate outcome; its content continues to inform and guide the agency as it formulates its actions.
Cell III – Complex law, wholly within principal-agent framework

III.A. Congressional dominance hypothesis: laws are often incomplete, but legislative coalitions passing statutes nevertheless get almost precisely what they want out of the bureaucracy through a combination of political controls (such as budget setting, oversight hearings, influence in the nomination process, and the ability to pass new laws) and legal administrative processes (such as strategically assigned burdens of proof, periods of delay, and requirements that certain viewpoints be considered). The seminal works in this tradition (McCubbins and Schwartz 1984; McCubbins, Noll, and Weingast 1987, 1989) argue that these factors systematically “stack the deck” in favor of preferred interest group allies when the law is interpreted and applied in the future. This perspective offers a subtle view of the role of law through its recognition that practical meaning will be shaped through ongoing debates about regulation and the settling of disputes through litigation. In this aspect, the work of McCubbins, Noll, and Weingast (1987, 1989) is a major source of inspiration for this dissertation, which repeatedly sounds these same themes.

In doing so, however, it abandons the strong-form “congressional dominance” hypothesis for a much weaker “statutory importance” one. The strong version shares with the simple models above a deductive insistence that the legislative principal (i.e., the “enacting coalition”) is able to engineer the law so it is capable of delivering just those policy results it wants (in this case, those that are amenable to its interest group allies). In this view, there is a complicated, intricate process of legal interpretation, but Congress, like a playwright, ultimately “dominates” this (basically determinate) process by scripting it in advance; details may be improvised, but the broad outlines are set. Empirically, this is dubious; evidence suggests that Congress is likely to fail in its attempts to privilege specific interests (Balla 1998). Theoretically, although these
works recognize the importance of the legal battles fought in bureaucracies and courts, they fail to provide much insight into how the impact of law will vary based on its substance. If statutes have differential impacts on policy (rather than simply predetermining who controls the outcomes), we must closely examine the mechanisms through which the words of statutes are translated into meaningful constraints. In this way, we can make realistic assessments of legal efficacy.

III.B. Mainstream legal academic views: judges have various interpretive techniques (discussed at length in Chapter Two) for coping with ambiguous commands from principals, including textualism, intentionalism, and purposivism. Intense debates surround the legitimacy, desirability, and usefulness of these techniques, but from a positive standpoint there is little doubt that each is practiced. Defenders of each technique (from a normative perspective) admit that individual judges might distort the technique to serve their own purposes, but insist that although abuses are real, their way of thinking of law shows how it can be principled and effectively constraining. Even when criticizing rival techniques, most legal academics strongly reject the idea that disagreements about legal intricacies are merely clothing for an otherwise naked exercise of power (as those subscribing to the radical indeterminacy described in IV.C, below, argue). In spite of ambiguities and difficulties, bureaucrats and judges interpreting the law are best thought of (both positively and normatively) as attempting to be faithful agents to their legislative principals.

At many points in this dissertation, I engage the debates between textualists and their opponents in some detail. Oversimplifying, the heart of these debates is whether the principal’s wishes are honored best by prioritizing the letter or the spirit of the law. Intentionalists and
purposivists are far more likely to imagine a legislative body capable of exerting a singular will like an individual, while textualists doubt that this projection corresponds to any underlying reality and therefore argue for restricting interpretive inquiries to what the text contains. Despite their significant disagreements, these approaches share a basic orientation: both assume that the legal interpreter’s task is to remain faithful to the principal’s commands as embodied in the law (Manning 2005, 421).

Why the Cell III models stop short of realism – beyond the principal-agent foundation

Although the models in Cell III treat statutory interpretation as complex, there is another level of complexity or realism to account for, which is that judges and bureaucrats only sometimes see themselves as attempting to act as faithful agents; other times, they see themselves as guardians of legal coherence, workability, and justice, muddling through where the law provides them little guidance.

21 In discussing the common ground between textualists and “classical intentionalists,” Manning (2005, 445) notes: “In any system predicated on legislative supremacy, a faithful agent will of course seek the legislature’s intended meaning in some sense, and modern textualists do situate themselves in that tradition.” Although modern textualists have famously questioned the existence of any “real” legislative intent, then, Manning shows that this does not negate their allegiance to a principal-agent way of thinking. Instead, it makes them skeptical of anthropomorphizing the legislature so to better understand its intentions by treating outside materials as evidence of a “thought-process.” To the textualist, an inference about a statute’s real meaning based on a single legislator’s remark in a committee fails to take into account the “cumbersome, chaotic, path-dependent, and opaque character of the legislative process,” and consequently textualists argue that the meaning of the statute should be understood wholly from asking how a reasonable reader would understand the final text. I argue that it is often unreasonable to treat a problem as having been addressed by the legislature at all, because neither textual meaning nor other indicia of legislative intent provides determinate guidance.

22 Throughout this dissertation, I have adopted the terminology of “muddling through” to describe this process. I find that this theoretical perspective (loose as it may be) helps to capture what policymakers do when “just following the law” stops providing them useful guidance.

What does it mean to muddle through in the context of interpreting a law? Lindblom’s (1959) classic formulation of “muddling through” thinks of a policymaker trying to solve problems, with the law hardly making an appearance in his discussion. When an official feels he has a clear and binding legal obligation, most often there simply is no problem; the appropriate course of action is just to follow the law. Where legal ambiguity creeps in or practical difficulties with following the legal path of least resistance arise, however, the law now becomes an element of the problem to solve, rather than a trumping solution.

As Lindblom argues, there are two basic ways of approaching complex choice problems. The first is synoptic: comprehensively accounting for all objectives, values, theories, and facts relevant to the situation at the
At least in the realm of statutory law, there is a strong (but not indefeasible) presumption that government officials in the executive and judicial branches should act as faithful agents by being obedient to the rules duly enacted by the constitutional legislative process. There are powerful norms of democratic legitimacy based on a theory about how the popular will can be transmitted through democratic representatives deliberating as a body, and this theory is reflected in the Constitution’s structure.\(^2\)

This fealty to the congressional principal does not trump all other values, however, nor is it best understood as providing useful guidance in every situation. Before they can conceive of themselves as rule-followers in any given context, bureaucrats and judges must answer two questions in the affirmative: 1) Has the legislature actually provided guidance sufficient to answer the interpretive question at hand? 2) If so, are the consequences of following its

---

\(^2\) This logic has clear limits, though, as the Supremacy Clause of Article VI makes clear. Since the Executive and Judiciary are ultimately constitutional bodies rather than mere instruments of the legislature, their most fundamental loyalty must be to the requirements of the Constitution; their constitutional status would be of little worth if they were willing to accept as their primary responsibility the transmission of legislative will, constitutional or not (see Paulsen 2003). One might nevertheless think that legislative enactments consistent with the Constitution should be the touchstone of all of the government’s actions, at least within the realm of domestic regulation that is the focus of this dissertation.
instructions acceptable in terms of reasonableness, coherence, workability, and justice? (Schauer 2009, chap. 8).

If the first question can be answered in the affirmative but the second cannot, then some exception to the controlling text may be found. In the constitutional realm, this wisdom is often sloganized as, “The Constitution is not a suicide pact”; in the statutory realm, we might say that no law is worth committing gross injustices for. Chapter Two, Section 3.C discusses exceptions to following clear text.

My main concern is with what answering the first question negatively entails.\(^{24}\) If the legislature is perfectly capable of providing determinate guidance but has not done so in a way that resolves the question at hand, the actors must still choose how to proceed. At this point, I argue that judges and bureaucrats will adopt a more prospective and independent attitude, taking into account various non-legal factors such as institutional capacities and practical policy considerations, rather than becoming preoccupied with the idea of faithful agency and trying to tease out what performing this role would require of them.\(^{25}\)

As the discussion above of Cell II suggested, one could simply say that where legal interpretation fails to provide determinate guidance, a distinct and non-legal activity called “policymaking” begins. I find this implausible, and instead argue that even when law is indeterminate it continues to play an informing and constraining role. In this regard, I make an analogous distinction to the one Whittington (1999) makes between constitutional interpretation

---

\(^{24}\) One might, instead, completely reject the premise of the question and argue that even for statute law, bureaucrats and judges are not primarily obligated to obey the will of Congress. See the discussion, below, regarding IV.A.

\(^{25}\) Political scientists in the Judicial Politics tradition have emphasized different non-legal factors, with “attitudinalists” stridently arguing for the importance of ideology and political party to the exclusion of others. More recently, proponents of the “strategic” view additionally emphasize the potential for political constraints on judicial pursuit of their own ideology through various mechanisms of punishing (or threatening to punish) judges, e.g., Clark (2010). I am sure that these factors are sometimes important—especially for highly salient and polarizing issues (such as “culture war” issues in recent years). Throughout this dissertation, however, I try to say why a focus on these factors is misleading in the context of complex regulatory policy questions, where I find both legal values and pragmatic policy considerations to be more important.
and constitutional construction (see Chapter Two). I argue that the substance of the law is still very much an influence on decisions even where it fails to provide determinate guidance. Whether a statute allows an agency administrator to take some action “if necessary,” “if, in the Administrator’s judgment, it is necessary,” or “as the Administrator deems reasonable,” it is clearly not providing exhaustive guidance about how to act; at the same time, it seems like a mistake to assume that the agency is given an equivalent, law-free opportunity to “make policy” by each of these common statutory formulations.

Instead, the practical effects of legal language must be learned through empirical investigation. Proceeding from an ambiguous, indeterminate text (construction) is certainly qualitatively distinct from deriving a single answer from a determinate text—but it is not likely to match the actor’s preferred approach if they were starting “from scratch,” with no law bearing on the question at all.

But why do such situations arise? Why does law fail to provide determinate guidance? Legislators have little incentive to push for completeness, and many reasons not to. It is often easier to agree on incomplete, open-ended statutes than to hash out a compromise on every point (Sunstein 1995). Even where legislators do resolve to pursue a clear goal, language is inherently difficult, the world extremely complicated, and the future impossible to predict. Just as contracting parties often rely on incomplete agreements because of bounded rationality and the costs of complete specification (Simon 1951; Williamson 1985), Congress often charges the Executive and Judiciary with making sense of incomplete statutes.²⁶

Thinking of legislation in this way takes us away from a notion of “control” that is pervasive throughout the political science literature: preference congruence, which states that if the actions of the agent (the bureaucrats) line up with the preferences of Congress, this shows

²⁶ For a classic account of how they did so with the Sherman Antitrust Act of 1890, see Letwin (1965).
effective control, while increasing divergence between bureaucrats’ or judges’ actions and legislators’ preferences shows shirking. Abandoning the simple models’ strong assumptions about legal transparency creates problems for this idea of control, but even more fundamental questions arise if we stop assuming that each institutional actor even thinks about or wants to control the answer to every practical question of policy. Because people tend to invest in acquiring only that knowledge they believe will provide net benefits to them (Hardin 2000), legislators may not even know enough about practical issues that arise under the statutes they pass to meaningfully be thought of as having “ideal points.” Even in those cases where they do learn of the question and have some preferred answer, if their constituents are likely to remain even more ignorant then gaining or losing “control” in the sense of congruent policies may not be a priority.  

27 As Arnold (1990, chaps. 3 and 4) explains, latent preferences can be powerfully activated given the right stimuli, but policy effects stemming from a long and complex web of actions and legal interpretations are rarely “traceable” enough to trigger these concerns effectively.

Law professors are more likely to recognize that statutes may not provide exhaustive guidance for bureaucrats or judges, but sometimes in narrow debates about proper methods of statutory interpretation they act as if these positions (all of which assume the rightness of a principal-agent framework) cover the whole field. Such debates are often interesting, but they

---

27 Approaching the point from a different angle teaches a similar lesson. Instead of asking, “How faithfully do bureaucrats follow the law?”, we can instead ask: “Where do bureaucratic goals come from, and what leads to the acceptance or rejection of these goals?” Empirical work suggests that goals only sometimes come from Congress. Just as often, agencies have to define their own goals. Wilson (1989) and Carpenter (2001, 2010) both vividly illustrate that one of the most important elements in determining the fate of an agency is how well its executives and “mezzo-level” permanent staff—potential bureaucratic entrepreneurs, in Carpenter’s terminology—can define a mission that is both appealing to broader political audiences and realistic given the tools available to them. Their actions will gain acceptance if the agency has or develops a reputation for competence. Heclo (1977) explains how difficult this task is likely to be when political appointees with only transient stints in government must try to gain the trust of the permanent staff. Statutes play little role in any of these authors’ work, which is in some ways problematic, but in any case shows that thinking of bureaucratic agencies solely as statute-followers (or disobeyers) is inadequate.
also have the sense of the eternal; it is difficult to imagine either side ever persuading the other to relent (and, as the conclusion of Chapter Two suggests, real practitioners have little reason to declare their fealty to one of these methods for all times and places).

My position is that there are two more fruitful questions to ask, each with a positive and normative dimension. 1) How readily, and in what situations, do (should) bureaucrats and judges stop thinking of themselves primarily as rule-followers? 2) What do (should) they do at that point? These questions are sometimes recognized and addressed\(^\text{28}\), but far less frequently than political scientists’ “who ultimately has control” or law professors’ “what is the right way to interpret statutes.”

One reason for this, alluded to in the discussion of Cell I, is that vulgar formalists and simple modelers alike have an interest in marginalizing the significance of the questions I present. Popular audiences often greet the suggestion that bureaucrats or judges should “actively” pursue values other than those set by formally enacted laws with cries of “bureaucratic overreach” or (implicitly illegitimate) “judicial activism.”\(^\text{29}\) As a result, few practitioners are able to speak frankly about these dynamics for fear of blowback. Meanwhile, a large elite part of the American Politics branch of the political science profession is heavily invested in their technology of spatial modeling, so that fundamental challenges to this approach are treated as

\(^{28}\) One expansive literature that addresses these questions, sometimes elliptically, is the philosophical jurisprudence literature, which seeks to understand exactly what law \textit{is}. Because this literature concerns what makes laws legitimate and authoritative and I do not take up these questions in this dissertation, I do not engage with it. Thinking at a deeper level about why laws ever command authority would require going down this rabbit hole, a task I hope to take up in future work. Just a few seminal works that demand attention are: Hart (1958, 1961), Fuller (1958, 1969), Raz (1990), and Schauer (1991).

\(^{29}\) As I note elsewhere, following Smith (1999), people thinking about the law in this way could arguably have salutary consequences for the health of a legal system. Nevertheless, from an analytical point of view these positions are quite unhelpful.
unwelcome, or somebody else’s problem. Needless to say, neither of these factors presents a compelling reason to value the line of inquiry suggested here any less.

Cell IV – Complex law, beyond principal-agent

The theoretical frameworks in this cell are the most realistic (which, again, does not imply they are the most suitable for every analytic purpose). They recognize that legal interpretation is complex and contested. They also realize that thinking about actions under the law entirely in a principal-agent framework is insufficient. The first grouping discussed (IV.A-IV.C) actually goes even farther, suggesting that the principal-agent framework is always a misleading and impoverished way to think about judging. The second grouping (IV.D-IV.F), which includes the position I build and defend in this dissertation, embraces the principal-agent framework as generally applicable but then goes on to discuss what happens (and should happen) when the law fails to provide determinate guidance.

IV.A. Judicial responsibility to justice: although ambiguity is frequent and leads to many hard cases, complications are in principle soluble, meaning that there are fundamentally right answers to every interpretive question. These interpretations are accessible to human reasoning, although by no means are these contained entirely within the words of a relevant constitution or statute. This is the “model” that Dworkin’s Hercules represents (see Dworkin 1978, chap. 4). Adopting this outlook is not meant to deny that getting the right answer is difficult—indeed, Dworkin’s terminology suggests that getting “fit” and “justification” correct is

---

30 This assertion is bound to come across as ungracious, but I believe it is important for people to make such statements to combat disciplinary pressures. In their plea for more multi-method research, Poteete, Janssen, and Ostrom (2010) note that career incentives serve as important methodological constraints, especially on young scholars.
a Herculean task, in which the interpreter must struggle to correctly balance society’s deepest moral commitments. Legal determinacy, in this light, is less a matter of practical likelihood (for real people are all too human) but of metaphysical theorizing. Dworkin’s scholarly ambition is not to provide empirically-based or institutionally contextualized explanations predictions about the practical effects of law, and therefore it would be unfair to criticize him for failing to do so.\(^3\)

Since that is the scholarly ambition of this dissertation, however, Dworkin’s philosophical concerns about the nature of law are put to the side. What is most notable about his view in this context is that Dworkin’s approach suggests that judges should not think of themselves as mere agents of the legislature; rather, they must “fit” their interpretations into a far larger web of legal meaning, utilizing materials well beyond the pronouncements of the legislature, including rather murky ones such as the society’s moral traditions and deep-rooted political values.

Dworkin is far from the only one to see judges’ primary fealty as being to something other than a particular statute passed by the legislature. Many proponents of natural law follow the maxim *lex inusta non est lex* (an unjust law is not law). Since all government officials owe their highest obligation to natural law, they should only follow statutory law if it is consistent with natural law imperatives (George 2008). Taking a somewhat different tack, there are those who argue that judges ought to be guided by “the equity of the statute” in resolving cases, explicitly privileging values of equitable judgment over fidelity to the plain text of the statute. In doing so, judges should extend statutes to encompass cases within their understanding of a statute’s purpose, and trim them to avoid harsh outcomes that do not seem to serve the purpose (see Manning 2001 for a discussion and critique of this position).

\(^3\) Posner (2003, p. 955) raises this defense of Dworkin against attacks that Dworkin is blind to empirical facts about institutional capacities, noting that Dworkin may still be criticized fairly for his many “confident evaluations of particular interpretive issues” without having done the necessary work to support these conclusions.
Although it is much more common to think about judges in these terms than bureaucrats, Shapiro (1988, 150-56) explains that there are those who believe that the administrative process leads to right answers, all-things-considered. In this view, bureaucrats’ deliberations about a statute’s true meaning ultimately lead to a true reckoning of the public interest; the statute’s requirements, rightly understood, go well beyond any questions of what the enacting Congress said or intended. Shapiro searchingly critiques the workability of this theoretical posture, which he believes leads to a focus on formal appearances to the detriment of workable substance.

IV.B. Legal realism (skeptical version): because legal rules collectively under- (or over-) determine legal meaning in most cases, judges end up deciding based on situation-types and their sense of existing norms, with legal reasons applied as post hoc rationalizations rather than being true causes of the decisions. Judicial behavior is predictable, stable, and even often determinate—but not a function of statutory language or any other formal rules, which are quite inadequate to the task. Instead, judges are guided by less formal and harder-to-articulate norms about the appropriate role of the judge. The idea that judges view themselves as nothing more than agents for the legislative principal is naïve. Leiter (2007, chaps. 1 and 2) attributes this position to the Legal Realists in his philosophical reconstruction of their works. Interestingly, the interpretive tasks faced by executive branch actors are almost never considered in the same

32 Twining (1973, 81) notes that Llewellyn was the one scholar ever to accept “the title ‘realist’ without reservations,” and warns that making categorical generalizations about “the realists” is probably hopeless as a historical matter. My own (limited and Llewellyn-centric) reading of the Realists leads me to believe that they would have admitted a greater possibility for determinate law (at least in their less-polemical moods) in the statutory realm. See the extensive references to Karl Llewellyn’s works in Chapter Two for substantiation of this claim. Llewellyn’s (1950, 399) famous comments about the canons of statutory interpretation admittedly seem to suggest the opposite. The canons and counter-canons he identifies give the judge flexibility to render a decision in either direction, so that judges can reach the results that strike them as sensible by selectively sampling from the set of “mutually contradictory correct rules.” But Llewellyn’s focus in the piece is on the “hard cases” found in appellate decisions, rather than on the overall potential of law. In his posthumous Theory of Rules (2011), Llewellyn emphasizes the possibility of clear rules creating binding meanings.
breath as the legal realists, whose central focus was on appellate judging. One could, however, imagine bureaucrats with a similar mindset quite easily. Faced with numerous administrative tasks and practical problems to solve, bureaucrats could evolve consistent and context-sensitive responses independent of any statutory language, only later finding convenient statutory language to justify their actions. To summarize this view, government officials are nearly always capable of providing legal justifications for their preferred actions. As a result, they take the actions they believe help them to pragmatically seek their goals, which derive from informal norms about their institutional roles.33

IV.C. Meaning-skeptical cynicism: because the meaning of words is always indeterminate, legal interpretations are always best thought of as exercises of political power, most often privileging the interests of the powerful. This is roughly the view of the Critical Legal Studies movement. This view is not equivalent to the simple-cynical model of I.C, since it does not assume that there are readily apparent right answers to legal questions that government officials choose to disregard. Instead, its “critical” posture goes deeper than the simple cynic’s: law’s pliability is endemic, rather than merely a consequence of agent willfulness. Still, while their conceptualization of the judging differs from the simple-cynical view of I.C, the deep skepticism about the meaningfulness of law that they ultimately adopt is similar. For anyone holding this view, the present project is clearly quite ill-conceived. Chapter Two describes why this view sets its few adherents apart from nearly everyone else who thinks about the law.

33 See Heclo (2008) and Murray (2002) for informative discussions of role-based norms. Precisely defining the content of these roles is very difficult, and scholars show a troubling lack of humility when they assume that the role, properly understood, is really quite simple, such that deviations from their idea of it should simply be written off as opportunism or stupidity. Whittington (2000, 621-25) similarly criticizes those who reduce law to a pure instrumentality and gives a brief review of recent works exploring the constitutive and role-structuring nature of law.
IV.D. Textualism + deference to executive branch: a statute’s text should be controlling where it is clear; where it is ambiguous, any reasonable construction by the executive branch should receive judicial deference. As Chapter Two explains, this logic closely tracks the two-step rule set by *Chevron v. NRDC* (1984), the case which for a quarter-century has cast a lengthening shadow over the practice of administrative law. Many prominent legal academics have also come to embrace the idea that judges should assert themselves as agents of the legislature in enforcing clear textual requirements, but should largely withdraw at that point and defer to the superior expertise of administrative agencies. For instance, Sunstein (1998, 1019) argues, “[W]ithout much fanfare, agencies have become modern America’s common law courts, and properly so,” approvingly citing their institutional capacity to work out the practical meanings of open-ended statutes with a sensitivity for real-world consequences.

Vermeule (2006) presents one of the strongest (both well-argued and uncompromising) versions of this argument, and so is worth considering at some length. Vermeule argues for unflinching adherence to clear textual meaning joined with judicial self-abnegation in the presence of ambiguity on the grounds that the judiciary’s institutional capacities make it ill-suited to pursue any more ambitious or active course (such as adopting a “democracy-promotion” rule). In Vermeule’s estimation, if judges go looking for exceptions to clear text or attempt to determine the best interpretation of ambiguous text, they are likely to incur significant costs to acquire the necessary information. At the same time, their “casual empiricism” is likely to lead them into error, and any social gains from well-conceived judicial interventions are likely to be swamped by losses from misadventures. On the other hand, if judges confine their inquiry to the text’s clear requirements, their costs will be low, their results predictable for potential litigants, and responsibility for error correction will fall clearly on the legislature. Agencies, on
the other hand, are intimately acquainted with operational details and the nuances of meaning in
the statutes they administer, such that they are likely to get things right when they diverge from a
strict formalist path (chap. 7). Vermeule’s position is driven primarily by system-level
institutional considerations; he argues that his position best accounts for what we know about
institutional capacities. Those arguing that judges can, do, and should take into account more
case-specific factors end up with “a chronic, systematic tendency to overestimate judicial
capacities, to underestimate the capacities of legislatures and agencies, and to ignore the
systemic costs of judicial discretion for other institutions” (59). Notably, this position’s support
for judicial textualism derives from a completely different source than the “rules is rules”
position of I.B (vulgar formalism), favoring an empirical assessment of institutional tendencies
rather than a formal statement of institutional competencies.

IV.E. **Weak textualism + judicial and legislative updating**: text should be controlling
if it is both clear and practically unproblematic, but as laws age and erode, judges should be able
to update them through changes in interpretation. This roughly describes the “dynamic”
approach to statutory interpretation favored by Eskridge (1994), who argues that as conditions
change the idea of fidelity to a statute’s enactors becomes chimerical and distracting. Any theory
which required static meaning would paralyze real-world interpreters, whose first priority is
resolving practical problems. Calabresi (1982) stakes out a similar position, drawing special
attention to the problem of “statutory obsolescence.” He argues that judges appropriately act as
common law updaters, but laments that they often do so through strained applications of
constitutional doctrines (chap. 2). Instead, judges should thoughtfully allocate the burden of
inertia through (as-yet-hypothetical) tools such as forced sunsetting of problematic laws that
would force legislative reconsideration. Once again, these authors emphasize institutional capabilities in arguing for their fitness as interpreters of less-than-limpid statutory requirements. Unlike Vermeule, they favor relying on judicial expertise and the legislature’s democratic legitimacy over the bureaucracy’s superior local knowledge.

**IV.F. Weak textualism + institutionally sensitive muddling through:** once again, text should be controlling if it is clear and practically unproblematic. If it is not, though, government actors attempting to resolve problems within a statutory scheme will eschew any hard and fast decision rules, instead favoring a context-sensitive assessment of institutional capabilities to guide their choices. In other words, bureaucrats and judges consider it a first-best option to act as the legislature’s faithful agents in effecting well-designed, coherent, and workable statutory requirements. At the same time, practical implementation problems, statutory lacunae, and ambiguous language confront them with many situations in which conceiving of themselves as mere rule-followers will either lead to results they consider unacceptable or leave them without meaningful guidance. At this point, they will muddle through as best they can with the information available to them, including about policy-specific institutional capacities. This dissertation mostly presents a justification for this account.

In many situations, this approach will lead to actions equivalent to those recommended by Vermeule in IV.D. When a statute’s text is clear, both bureaucrats and judges will hold themselves to its requirements. When there are subsidiary questions that arise in the course of administering the statute that Congress seems never to have explicitly considered, it makes eminent sense to allow the agency charged with implementing the law to make reasonable choices to fill those gaps.
In other situations, however, the interpretation of key statutory phrases may effectively determine the scope of the policy, as well as its ability to keep pace with changing conditions. Determining how these questions are answered will have deep political consequences (in terms of winners and losers) and policy consequences (in terms of coherence). While superior executive-branch expertise is quite plausible for minor practical issues, it is far less likely to be relevant for fundamental questions—especially those that clearly implicate a large number of policies administered by other agencies. System-wide coherence is unlikely to be the priority of any single agency, and in some circumstances judges may justly believe themselves to be better positioned to maintain coherence than bureaucrats.

Another relevant factor is Congress’s practical ability to correct problems caused by obedience to the statutory text. As Chapters Four and Five argue, textualism rests on the premise that the legislature can, must, and will respond to deficiencies in the law by amending the legislative text. But Congress’s monitoring ability and will are inconstant, and the likelihood of its acting on behalf of coherence often looks dim. Contra Vermeule, most judges do not consider the assumption that they are unable to effectively understand such context to be a desirable form of institutional modesty. For instance, Judge John Newman, then a member of the U.S. Court of Appeals for the Second Circuit, wrote (1984, 209-210) that he favored strict textualism for a law like the Internal Revenue Code, which Congress amends on a yearly basis, but that statutes such as 42 U.S.C. §1983, enacted in grand historical moments and unlikely to be revisited, make a strict adherence to a single interpretive meaning less desirable.

More generally, Newman argues that judges do and legitimately should rely on certain “institutional values”: “values that concern the judge’s conception of the role of the courts and other sources of law, the judge’s view of the federal system, and the judge’s informed sense of
the process of adjudication” (208). This guiding principle of interpretation has frequently been associated with Hart and Sacks (1994 [1958]), whose course materials on *The Legal Process* stressed the importance of accounting for institutional capabilities in deciding difficult legal questions. Perhaps unfairly, the central lesson scholars tend to have remembered from the legal process tradition is an across-the-board prescription for using courts to decide “dyadic” disputes but leave more complex “polycentric” issues to legislatures or administrators to resolve (see Barnes 2009, citing Fuller 1978). But, as Hart and Sacks argued, judgments about institutional capacities need not be made for all places and all times, or even in terms of whole institutions. Instead, policymakers attempt to incorporate the real-life institutional capacities for the specific problem they face into their thinking on how best to proceed. Instead of “is this a question best addressed by legislative deliberation, administrative processes, or judicial disposition,” a judge will ask “is this the sort of situation in which the FDA’s institutional advantages mean that its reading of the statute deserves my deference?” (as in Chapter Four). This is the question that can more readily be answered, in spite of Vermeule’s doubts about judicial capacity.³⁴

The approach described here, with its prominent place for muddling through, has one clear disadvantage relative to the model of judging outlined in IV.D which couples strong textualism with thoroughgoing deference to the executive in the presence of statutory ambiguity: saying that final decision-making authority will change depending on the circumstances makes it harder for government officials and regulated parties alike to predict how and when policy will

---

³⁴ Sunstein (1998) and Vermeule (2006) are two of the most powerful statements arguing for a more-or-less context-insensitive regime of deference to executive choices when statutes are indeterminate. In other writings, though, these same authors seem to take the opposite position, recognizing the need for context-sensitive evaluations—see Sunstein (1999) and Sunstein and Vermeule (2003, 886), which argue that the most important question to answer is “how should certain institutions, with their distinctive abilities and limitations, interpret certain texts?” In making this recommendation, they rather oddly suggest that no other scholars of legal interpretation have emphasized institutional capacities—a claim for which Posner (2003) takes them to task. Like the argument I advance here, Posner’s preferred brand of judicial pragmatism places great weight on the particular institutional facts of a case in determining what course of judicial action is appropriate.
ultimately be settled in an area. An exception-free textualism coupled with deference to the executive branch, on the other hand, offers permanent resolution of statutory meaning either at the time of a statute’s passage or when a regulatory agency decides how to construe some ambiguous text. According to Vermeule (2006), such a prospect should discipline the legislature both to take care with its statutory language and to correct errors actively rather than relying on judicial clean-up. This argument’s deductive logic is quite convincing, but whether this logic tracks actual institutional behavior is highly questionable.\(^{35}\) Certainly, judges often seem to doubt it. Chapter Five presents a highly problematic case study, where courts respond to growing strain on a statute by saying “this problem is Congress’s to fix,” only to see Congress spin its wheels in spite of its awareness and concern. Judges’ “casual empiricism” and a case study in this dissertation are hardly sufficient to resolve this question, but Vermeule’s work provides no systematic empirical evidence to resolve it in the opposite direction.

**Methodology and Goals**

What is the purpose of our journey from the simple models of Cell I to the complex, realistic, and murky world of my preferred model in IV.F? Why should a political scientist get bogged down in the details that muddling through says are important, rather than finding creative and novel ways of cutting through that complexity with general insights? For many research questions, this abstracting approach might well be inappropriate; each of the models I have

---

\(^{35}\) Arguably, “it won’t work until we really try”—that is, unless all judges stop cleaning up legislative messes, the legislators are unlikely to get a clear message about their own responsibility. Like a child who creates a mess, is told to keep his own room tidy, but finds that all of the toys magically return to their designated places when the mess becomes bad enough, Congress needs clearer lessons before it can learn. But this would be an ironic concession for Vermeule, who attacks interpretive programs other than textualism for being impossible to realize unless there was a perfectly coordinated judiciary. Having recourse to the argument noted here, though, suggests that textualism is just as ambitious and in need of near-perfect coordination, undercutting Vermeule’s claims for its modesty as an approach to judging.
discussed here might well be useful for various purposes.\textsuperscript{36} To understand what laws actually do in terms of guiding and constraining the development of policy, though, we need a realistic approach instead of beginning with simplifying assumptions about law’s transparency or the ways in which law is designed to further the goals of the legislative principal. In the following paragraphs, I describe this dissertation’s methodological techniques for pursuing this realism.

\textit{Descriptive}

The in-depth case-studies that largely make up this dissertation are, in some substantial part, descriptive. Although in some circles of contemporary political science “descriptive” has become a pejorative meaning “unworthy of academic attention,” descriptive work is a necessary part of advancing knowledge and should be accorded greater respect.\textsuperscript{37}

If we evaluate the usefulness of descriptive works by the same criteria we evaluate the rigor of large-n studies, the former inevitably look inadequate. Although it is not true that case studies are “n=1” research designs, since they employ within-unit diachronic variation (Gerring 2004, 344; Bennet and Elman 2006, 461), they nevertheless are less powerful for illuminating “true causal effects” that are the gold standard for much of the field. But if case studies are less capable of answering “why” questions, they have a distinct advantage in speaking to “what” happens and “how”; put another way, they are invaluable for illuminating causal mechanisms (Gerring 2004, 347-9; Bennet and Elman 2006, 457; Mahoney and Goertz 2006). Since trading

\textsuperscript{36}I wish to stress that my aim is not to wage methodological warfare—I agree with Ostrom (2006) that scholarly factionalism is a fruitless distraction and that using multiple methods is the most reliable path to understanding. Rather, I seek to understand what purposes these models can and cannot usefully serve (see Clarke and Primo 2011), thus enabling myself and future scholars who share my interests in the effects of the law to open other avenues of research where appropriate.

\textsuperscript{37}Of course, in other parts of the profession, descriptive work is alive and well. Relevant to this dissertation, members of the Law and Society community often seek to study the real effects of law for regular citizens, thus providing a “bottom-up” view of the law (e.g., Felstiner, Abel, and Sarat 1981; Greenhouse 1986; Ewick and Silbey 1998).
off depth and scope is inevitable, small-n qualitative research and large-n quantitative research are best thought of as complements.\textsuperscript{38}

Because my focus is on what and how questions, descriptive work provides a useful means of moving the project forward.

\textit{Agenda-Setting}

As their mathematical prowess has grown prodigiously in the last two decades, social science methodologists have tended to privilege the knowledge-value of confirmatory hypothesis-testing research over less structured, exploratory studies. Again, however, a broader conception of how knowledge expands finds a place for both types of work, and “[c]ase studies enjoy a natural advantage in research of an exploratory nature” (Gerring 2004, 349). Once again, there are tradeoffs; “the looseness of case study research is a boon to new conceptualizations just as it is a bane to falsification” (350).

This dissertation argues for the utility of such a “new conceptualization” in understanding statutory law’s role in constraining and shaping regulatory policy in America’s administrative state. As this Introduction’s taxonomy of existing models of statutory law illustrates, available theoretical frameworks tend to make radically simplifying assumptions about the nature of law, leaving them ill-suited to support a realistic view of what law does or to show how laws could be made better. These are not necessarily damning failures, as the research methods employed may fit the questions being asked. For my purpose of understanding what law does, however, a broader and less reductionist view is necessary. By choosing depth, I am able to establish the

\footnote{Leading statistical methodologists additionally urge caution in accepting the results of “black box” empirical studies which only seek to establish whether a causal effect exists without any investigation into the causal mechanisms showing how it functions. See Imai, Keele, Tingley, and Yamamoto (2011) and the sources cited therein for a review of these issues.}
need for greater attention to institutional contexts in understanding how laws are given practical effect through the processes of bureaucratic interpretation and litigated contestation.  

In doing so, I join an increasingly strong group of scholars who have rallied under the banners of “historical institutionalism” and “American Political Development” (APD) in recent years (see Orren and Skowronek 2004). This group has largely accepted the legal-realist-tinged lesson of the judicial politics and law and society literatures that “law in practice” often differs from law on the books (Whittington 2000, 620), but has sought to show how law and the judiciary must be understood within the rich institutional contexts in which they are embedded. In doing so, they have “rehabilitated elements of traditional public law” scholarship (613) and planted many seeds of inquiry in fields that had long lain fallow.

Similarly, this dissertation’s success must be measured in fruitful questions raised and avenues of exploration opened rather than in puzzles solved. At the conclusion of each chapter, I present empirical questions for further research. A major goal of my research is to show that answers to these questions would be of direct use to policymakers, who would be eager to utilize more systematic evidence about institutional capacities were it available to them. It is these questions’ novelty, relevance, and usefulness that justify the dissertation’s development of a

39 To be sure, there are other visions of how legal realism can be resuscitated in this golden age of large-n empirical studies; Miles and Sunstein (2008) argue that a “new legal realism” is currently flourishing in the legal academy in the form of large-n studies.

40 My work also joins a tradition of detail-oriented policy studies, which make theoretical contributions in large part by grappling with the processes of legal and regulatory change at the level of individual policies, e.g., Derthick (1979); Melnick (1983, 1994); Katzmann (1986); Rabkin (1989); Harris and Milkis (1996).

41 For instance, Whittington (2007) and Crowe (2012) have focused on how, why, and when the other branches of government sought to build up the judiciary’s institutional power. Frymer (2008, chap. 4) describes how changes to the rules of civil procedure in the 1930s and 40s allowed courts to carve out an independent sphere of power in civil rights litigation without any legislator intending this result, and thus provides support for this dissertation’s claims about law’s independent influence. Staszak (2010), in contrast, explores efforts to “reanchor” access to courts. Sometimes categorized as a part of APD and sometimes as a part of Law and Society, there is also a developing literature which emphasizes the importance of litigation to realizing legal change, e.g. Kagan (2001); Epp (1998, 2009); Barnes and Burke (2006); Teles (2008); Farhang (2010).
loose, descriptive conceptual framework rather than a tightly focused theory that delivers clear predictions.

**Beyond “Law vs. Politics”**

Ultimately, I believe this questioning spirit better honors the spirit of legal realism than do the long line of scholars who have devoted themselves to the dichotomous dispute of “law vs. politics” and, more importantly, is more likely to generate a research program that would deliver on legal realism’s promise of driving beneficial legal change. Taking Llewellyn (1930, 442-4) as my guide, the realists were never seeking to argue that rules of law are unimportant; rather, their argument was always that rules were only important inasmuch as they affected real behavior, a matter that required empirical investigation rather than blithe assertion. This hardly meant rules of law would recede from the discussion, though: “Whether they be pure paper rules, or are the accepted patter of the law officials, they remain present, and their presence remains an actuality—an actuality of importance—but an actuality whose precise importance, whose bearing and influence become clear [through inquiry]” (449). By showing how the law functions—really—this empirically informed perspective should be of great use to the practicing lawyer, the legislator, the philosopher of law, and the judge (446n12).

---

42 See Whittington (2000, 629) for a similar expression of frustration with the persistent fascination with this dichotomy.

43 As Twining (1973, 148) notes, two of Llewellyn’s most famous pronouncements from his most widely read work, *The Bramble Bush*, exhibit an arch-skepticism that did not really represent Llewellyn’s considered view. “What these [legal] officials do about disputes is, to my mind, the law itself,” and the memorable designation of rules as unimportant “except as pretty playthings,” were both sentiments that Llewellyn disowned in the book’s second edition, where he printed a retraction and expressed regret at his use of such “unhappy words.” As Chapter Two’s numerous citations to Llewellyn’s posthumous *Theory of Rules* (2011, originally written in the late 1930s) show, Llewellyn had few doubts that law could be binding.

44 That this is the adverbial form of “real” has been a constant frustration for me; endlessly repeating that you are interested in what is “really” going is a recipe for bad prose. “Actually” and “in practice” are not much more appealing. Nevertheless, these words really get at the point, and so I rely on them heavily.
Nevertheless, for many people the realists are remembered most for having “stressed the role of personal choice in judging” and thus contributed to the “founding debate of judicial politics—is Supreme Court decision making driven by law or politics? —[which] remains at center stage today” (Lax and Rader 2010, 273). More than just trying to establish that political factors are present and important, recent seminal works of the judicial politics literature (Segal and Spaeth 2002) have taken pains to rhetorically stress that law’s importance is negligible, and while these claims are very often aimed at just a small portion of the Supreme Court’s docket (e.g., Segal and Westerland 2011, which examines only judicial review cases) they seldom emphasize this limitation, and instead often disparage the uselessness of the “legal model.”

Although things in the field are changing, with many scholars seeking to understand law in creative new ways (e.g., Richards and Kritzer 2002; Clark and Lauderdale 2010), few have made the shift from investigating “whether” law’s causal effects can be decisively proven across a large body of cases to understanding “what, exactly” law might be doing in shaping the development of policy.

---

45 This view is generally supported by treating Jerome Frank as the representative realist, when in fact he was both intellectually and temperamentally an outlier; see Llewellyn (1931) and Leiter (2007).

46 It is worth noting that the attitudinal model does admit the importance of case facts to justices’ decisions, which sometimes led to analyses (e.g. Segal 1984) very much in the realist tradition. Unfortunately, as the attitudinal model aged and was increasingly contrasted with the “legal model,” its users tended to emphasize the role of judicial ideology more and more. There are promising signs of change, however; see Friedman and Martin (2011).

47 Noteworthy exceptions include high-quality empirical work coming out of legal academia in recent years, such as Gersen and O’Connell (2008), Eskridge and Baer (2008), and Raso and Eskridge (2010). Unfortunately, some of the most impressive large-n empirical studies to come out of legal academia in recent years have rather uncritically adopted the law vs. politics framing, e.g. Revesz (1997), Miles and Sunstein (2006), Sunstein and Miles (2008).
Although in the planning phases of this dissertation I hoped to pursue regulatory statutes’ causal effects through a large-n investigation, as I progressed I largely shifted my efforts toward deep engagement with the policy particulars of my cases. To some, this is the very definition of “getting bogged down in details,” but there are reasons to think that sensitivity to the particulars of law and policy are necessary to realize legal realism’s promise of making useful contributions to the functioning of the legal system. At several points, Llewellyn suggested that this would be the case, and he bears quoting:

In view of the tendency toward over-generalization in the past [progress in research] is likely to mean the making of smaller categories—which may be either sub-groupings inside the received categories, or may cut across them. (Llewellyn 1930, 453)

From another angle, what we need is patience to look and see what is there; and to do that we must become less ambitious as to how much we are going to look at all at once. (Llewellyn 1930, 457-58)

[We should look to make] the study of law a study in the first instance of particularized situations and what happens in or can be done about them. (Llewellyn 1930, 460)

[W]e can learn from natural science that Big Things in the pursuit of science are not commonly achieved by going after Big Things. They come, vastly more, out of sustained, insistent, cumulative digging after small bits of testable and tested knowledge about small things; and out of concentrated study—including speculation—on matters little enough to be studied closely. (Llewellyn 1941, 14)

With this advice in mind, we can see a major opportunity to improve our understanding of the law through detailed engagement with its particulars. If we are interested in how a single law functions, especially over a long period of time, we will necessarily have to grapple with a few technical breakthroughs (or acquire a small army of talented research assistants) to make this feasible.

---

I had hoped that the extensive data collected for Chapter Three would provide me a sufficient basis for matching the content of statutory text with the results in terms of bureaucratic constraints, but, alas, the data proved unequal to these purposes. My plan to match statutory subsections with bureaucratic actions proved more cumbersome than I was able to imagine: laws in action are vexingly difficult to cleanly subdivide, and it is nearly impossible to devise a mechanical procedure capable of accurately matching relevant textual requirements to specific government actions. Counting words or word frequencies is enticingly simple, but to date I have been unable to convince myself that these techniques are actually capable of delivering understanding of law’s effects. I remain committed to the project of scaling up this research and ultimately conducting large-n research, but will need to make some technical breakthroughs (or acquire a small army of talented research assistants) to make this feasible.
wealth of policy details. Understanding the policy as practitioners do (what Carpenter [2010, 29] calls achieving “dense knowledge”\(^{49}\)) will be necessary. Such a strategy creates significant challenges, both in researching a subject and in figuring out how to present findings transparently to those who have no expertise in the laws at issue. In general, however, acquiring a degree of expertise on a particular policy should be counted as a major benefit of a detail-oriented methodology, rather than a cost. Limiting scholarly inquiry to only those elements of a case study that are generalizable would “constitute a considerable waste of scholarly resources” (Gerring 2004, 345-6). The role of the citizen-scholar may not be entirely “scientific,” but it is no less potentially useful to society for that.\(^{50}\)

Llewellyn stands as an exemplar not only for what he said, but also for his actions. Later in his career, after decades of intensive study of commercial law, Llewellyn became one of the leading figures in the successful campaign to draft the Uniform Commercial Code that all of the states subsequently adopted, which simplified, clarified and regularized the nation’s commercial law. Though a huge number of lawyers and other interested parties were involved in the process of preparing the Code (and Llewellyn was not its most effective political advocate), the Code has sometimes been referred to as “Lex Llewellyn” or “Llewellyn’s Code” (Twining 1973, 271, 298). Llewellyn strove for a moderate position for the code: avoiding obsession with codification for its own sake and overconfidence in the drafters’ foresight, somewhat trusting of businesspeople and judges trying to make the system work, but nevertheless bringing well-crafted order to the system that would, on the whole, make things more predictable. The Code

\(^{49}\) Carpenter (2010, 29) lays out the “dense knowledge test” as follows: “Does a theoretical model generally, and an empirical account specifically, make sense to those most thoroughly and intimately aware of the action? Do quantitative analyses count up events that historians, ethnographers, and careful observers of the events would never consider comparable in the ‘apple to apple’ sense?”

\(^{50}\) Llewellyn’s (1941, 22) language is again felicitous: “Knowledge does not have to be purely scientific, in order to be on the way toward Science. Neither does it have to be scientific in order to be extremely useful.”
was drafted with the practical problems of interpretation in mind, offering explicit statements of purpose and policy goals, a mandatory rule of liberal construction. It was also accompanied by an extensive commentary by the drafters which further clarified the purposes through illustrations and discussions of particular cases (Twining 1973, chaps. 11 and 12). In short, Llewellyn orchestrated a process which applied the best available empirical knowledge about real legal practice to improve the system in ways that were both coherent and workable.

This is an example political scientists would do well to consider. I would like to say that the four chapters that follow make a down payment on the research agenda demanded, but in these post-crisis years I suspect that their contributions are not quite large enough to meet the underwriting standards. My hope is that they nevertheless point the way while making substantial contributions to live policy debates of real importance.

In the remainder of this Introduction, I briefly situate each of these chapters within the framework discussed above. I then conclude by considering lingering questions and the techniques needed to advance the research agenda further.

Chapter Two – Analogizing Statutes, Contracts, and Constitutions

The preceding discussion and most of this dissertation analyzes statutory interpretation, but much of the reasoning applies equally well to the interpretation of other legal documents, including private contracts and constitutions. Chapter Two seeks to derive important insights about the general practice of interpreting legal texts by pursuing the analogies between these three branches of law.
The position I stake out across each of these domains is the same theoretically modest one I sought to explicate in the preceding discussion: clear text must be controlling (though there are illuminating exceptions), or else the very idea of using binding textual instruments is useless; importantly, though, contracts are rarely crystal clear or complete. Applying Whittington’s (1999) distinction between constitutional interpretation and constitutional construction to all three areas, I argue that where text is indeterminate those charged with making sense of the text define its meaning through usage and contestation. Their acts of construction are rooted in, but not determined by, the content of the textual instrument and necessarily incorporate values not found within the text. In the language of this Introduction, bureaucrats and judges find that when textual meaning runs out, conceiving of themselves as faithful agents of the textual instrument’s drafters fails to provide answers, and so they must have recourse to non-textual criteria as they give the law practical meaning.

This far, I argue, nearly all students of interpretation are able to agree. Beyond this, however, there are real disagreements. Whose constructions of indeterminate text should be controlling? Although there seems to be a basic agreement that institutional competencies should matter, strong divergent prior beliefs about various institutions (as well as how much weight to give to formal designation of competencies) lead to heated exchanges. As in this Introduction, I argue that attempting to provide universal empirical answers to these questions is likely to be fruitless—both because doing so convincingly would be so daunting, and because actual policymakers are unlikely to trouble themselves with such abstract affairs. Instead, I explain how political scientists and other scholars should look for opportunities to answer questions about context-specific institutional capacities.
Chapter Three – The Grammar of Legal Determinacy, Contestation, and Constraint

Chapter Three argues that the law’s effectiveness as a constraint on bureaucratic judgment is inextricably linked to who can effectively challenge agency legal interpretations. The presence and efficacy of challenges, in turn, depend upon the determinacy of the language in the law. Laws of high and low determinacy lead to relatively few conflicts because the outcomes of those conflicts are fairly predictable. Medium determinacy laws—often those whose pivotal sections depend on the meaning of adjectives, such as “necessary” or “appropriate”—are more likely to engender conflicts, as the fact of legal constraint on the executive is clear, but the exact nature of that constraint is ambiguous and contestable. In the language of this Introduction, agents are effectively bound by very clear statutory requirements. Where a legislative principal has set a clear requirement, an agency will rightly expect to face a successfully litigated challenge if it fails to abide by that rule. On the other hand, where a legislative principal makes clear that an agency’s judgment is to be decisive, the statute will, as intended, allow bureaucrats to devise their own course. Conflicts are likely to arise when statutory text makes it clear that the legislative principal intended to create some constraint, but when the precise meaning of that constraint is unclear.

To illustrate this dynamic, Chapter Three provides the dissertation’s broadest portrait of statutory text being turned into practical requirements for regulated firms by focusing on Environmental Protection Agency (EPA) rulemaking under the Clean Air Act. An original large-n dataset is used to show the prevalence and importance of commenter participation and litigation challenging agency interpretations of the statute. Additionally, close examinations of several specific parts of the law provide examples of how statutory language’s determinacy shapes legal conflicts about a law’s practical meaning. Far from being transparent, as the models
of Cell I assume, we see that the legal requirements in the Clean Air Act are anchored in verbal formulae of uncertain bite, “adequate margin of safety,” “ample margin of safety,” and “maximum achievable control technology” being important examples. Litigation often focuses on just what interpretations these phrases can bear, as well as whether the EPA’s expertise justifies it in taking the law in unanticipated directions.

Chapter Four – When Can You Teach an Old Law New Tricks?

From the viewpoint of an omnipotent, generic policymaker, policy changes are most readily made through changes in statutory text. I argue, however, that for real policymakers situated at particular institutional vantage points (as well as for interest groups outside of government) seeking legislative change may not be the most attractive means of changing policy. Instead, a purported deficiency in government policy can be portrayed as an interpretive problem: the statutory text could support bold action, but the interpreters of the policy are failing to utilize the existing statutory tools. Such a claim can be asserted in court, leading to the possibility of changing policy without having to get a bill through Congress.

Chapter Four clearly sounds the dissertation’s themes when it considers how judges are likely to respond to “old law, new trick” maneuvers. I argue that, along with the familiar interpretive considerations of text, legislative intent, and deference to the executive branch, judges are also likely to look to institutional competencies as they decide whether to support a novel interpretation of an older statute. Using two detailed case studies—the Clinton FDA’s assertion of authority over tobacco using the Food, Drug, and Cosmetic Act, and environmentalists’ attempt to force the EPA to regulate greenhouse gasses under the Clean Air Act—I argue that judges must be convinced that the normal policymaking process has proven
inadequate before they are willing to support the second-best option of adapting an old statute. A Supreme Court majority found that Congress had consistently shown its ability to regulate tobacco, and therefore struck down the FDA’s novel interpretation in *FDA v. Brown & Williamson Tobacco Corp.* (2000). Conversely, a Supreme Court majority in *Massachusetts v. EPA* (2007) found that the Clean Air Act was indeed required to cover greenhouse gas emissions as a “pollutant,” in spite of the EPA’s objections, noting Congress’s fecklessness in responding to the dire threat of global warming. Although the judges in these cases concentrated much of their rhetorical firepower on the question of what the statutory text “clearly” requires, the chapter argues that the institutional perspective does a better job illuminating the real differences between the judges. The chapter also considers the long-term consequences of changing policy through new interpretations rather than new laws, showing that there are reasons to worry about the perverse constraints exerted by clear and binding statutory text in a statute that no longer serves the function its drafters envisioned.

**Chapter Five – Statutory Maintenance and Erosion: The Case of the Glass-Steagall Act**

Continuing to chronicle the lives of aging laws, Chapter Five seeks to understand roughly the mirror-image dynamic of Chapter Four: instead of seeing how laws’ scope can be expanded through interpretive entrepreneurs, I examine how a regulatory regime’s practical effects can be eroded. Through an in-depth historical case study of the Glass-Steagall Act and its separation of commercial and investment banking, the chapter shows how various legislative and judicial acts of statutory maintenance at first ensured that the law’s efficacy was not eroded by the exploitation of various statutory ambiguities and loopholes. The text’s tendency to be overwhelmed by changing business forms and legal creativity, with its efficacy dependent on
affirmative acts of maintenance by the principal, shows that statutory text sometimes requires 
dedicated and attentive legislative guardians to be effectively binding.

Beginning in the 1970s, however, the law faced an even more serious challenge when its 
bureaucratic administrators in the Office of the Comptroller of the Currency, Federal Reserve, 
and Federal Deposit Insurance Corporation began to deprioritize its goals. With commercial 
banks’ viability threatened by the rise of mutual funds, money markets, and other novel financial 
players, these regulators increasingly sought a workable policy that would strengthen banks by 
permitting their expansion into previously forbidden activities. Fairly clear statutory text stood 
as an obstacle to these efforts, and at first courts showed their willingness to act as faithful agents 
to even a long-gone Congress, clearly stating that it was the legislature’s job to modernize the 
regulatory regime rather than the courts’. But although Congress frequently considered reform 
legislation throughout the 1980s and early 1990s—and these policy changes even seemed to be 
broadly popular—no law passed until 1999. In the interim, the courts rather quietly withdrew 
their opposition to the bureaucrats’ consistent efforts to redefine the limits of commercial 
banking. Statutory text remained constant, but policy dramatically shifted as bureaucrats 
pursued values quite opposed to the intent of the departed principal. The judges’ behavior 
strongly suggests that they were willing to fight for a straightforward application of the statutory 
text, but only as long as the relevant institutional competencies made obedience to a legislative 
principal a reasonable course of action. When many years of legislative dysfunction called this 
approach’s soundness into question, judges were willing to accept bureaucratic adaptation as a 
second-best substitute. Even as the law’s efficacy waned, however, its clearest requirements 
continued to tie the hands of bureaucrats—often with perverse consequences.
Through this detailed case study, Chapter Five shows what factors are required for a the principal-agent dynamic to function successfully, as well as showing how this relationship can decay and ultimately fail.

**Lingering Questions and Future Research**

Among this dissertation’s most ambitious goals is to orient scholarly attention to questions susceptible to empirical research. As such, it raises more questions than it answers. Following in the footsteps of the Legal Realists, these questions center around *what law does*, rather than treating statutes as inert outputs of the political process. They include:

- What kinds of statutory text lead different institutional actors to be assertive? Are certain legal formulations especially constraining in practice?
- Do some textual formulations lead courts to overturn agency choices more often? What are they? Can this judicial empowerment be plausibly understood as intended by the legislature?
- What factors lead bureaucrats and judges to weigh textual requirements especially highly relative to other goals they may have? When are these actors willing to deprioritize text? To what extent do judgments about institutional capacities shape these decisions? What are the consequences when judges prioritize institutional capacities?
- When and why are exceptions made to clear textual requirements? What consequences does granting such exceptions have?
- Do government officials act as if they are bound as much by settled interpretations of ambiguous text as they are by clear statutory text? Does the manner in which the legal status quo was reached matter when addressing related interpretive questions?
- In any particular policy area, are bad (in any sense) policies usually directly attributable to clear legislative requirements or do they tend to be rooted in bureaucratic or judicial interpretations?
- Do laws work better (in any sense) when they create clear requirements, or is it actually efficient to work out ambiguous meanings through legal contestation?
- What is necessary to achieve legal coherence, workability, and reasonableness? In practice, do these values complement or compete with obedience to clear and predictable rules?
To answer these questions, we must do more than attend to laws’ legislative births. We must watch statutes mature, labor, and age in the hands of bureaucrats and lower courts. We must also be willing to engage policies at a very detailed level, a daunting and often discouraging task. But with enough information about what laws do in particular contexts, we might well learn enough about what works to be able to provide useful advice to lawmakers and others involved in the policymaking process. The next great research frontier for political scientists is to chronicle the lives of the laws, en masse.
Works Cited


Leiter, Brian. 2007. *Naturalizing Jurisprudence: Essays on American Legal Realism and


Mahoney, James, and Gary Goertz. 2006. “A Tale of Two Cultures: Contrasting Qualitative and Quantitative Research.” Political Analysis 14 (3): 227-249.


Chapter Two – Analogizing Statutes, Contracts, and Constitutions

The final three chapters of this dissertation seek to understand how statutory text constrains government officials through in-depth empirical examinations of the interpretation of particular statutes. First, though, I seek to better establish the nature of statutes, clearly articulating what kinds of social purposes they are meant to serve, how interpreters approach their text, and what their limitations may be. These questions are easier to address if we recognize that they are not distinctive to statutes. Indeed, the legal interpretation of all different kinds of textual instruments raise similar questions, and the interpretation of contracts and constitutions have each generated voluminous literatures that I draw on here.¹

In this chapter, I develop an extensive analogy between the social practices of interpreting contracts, interpreting statutes, and interpreting constitutions. Recognizing this analogy is hardly original—indeed, it probably occurs to anyone thinking about any one of these subjects who has even a passing interest in the others. In spite of this general awareness, the deep parallels across these different types of interpretation are not widely appreciated.² As a result, some of the keenest insights by scholars in each of these domains have not been applied to the others. Additionally, few scholars seek the lessons that might be gained from a more

¹ These are not the only types of textual instruments we could consider. Mahoney (2007) considers how international treaties ought to be analogized to statutes and contracts, arguing that principles from contract law capture the dynamics between treaty partners better than those of statutory interpretation.
² Legal scholars occasionally consider two or more of these types of interpretation together and consider whether they should be subject to the same kinds of thinking. Stack (2004, 2n1) provides a useful list of sources considering statutory and constitutional interpretation together and argues that justifications for formalism in one context do not necessarily transfer to the other.

Barnett (2004, chap. 4) offers perhaps the most widely noticed analogy between the interpretation of constitutions and contracts, arguing that the same considerations lead to a compelling case for textualism in constitutional law and contract law. I note those instances in which my argument closely resembles Barnett’s below.
comprehensive discussion of legal interpretation which situates contracts, statutes, and constitutions within a common frame of reference.

Beyond detailing these analogies—a useful organizational project in itself—this chapter attempts to contribute to the scholarship of all three areas of the law by showing how similar the theoretical battle lines are in each domain. Contract law formalists, statutory law textualists, and constitutional law originalists all insist that law can and should adhere to the commands of the text to settle nearly all disputes; neoclassicists, purposivists, and living constitutionalists insist that attention to far more the text is necessary to reasonably interpret these documents and give them power over our lives.

Though these clashes may to a large extent be permanent features of the legal landscape, a rapprochement is afoot in the realm of constitutional scholarship. Prominent and thoughtful adherents of warring camps have come together to declare that “We are all originalists now; we are all living constitutionalists now.” While some on each side think of this coming together as a deceptive mirage and denounce it as such (Alicea 2010; Koppelman 2011), I believe more optimistically that a real intellectual synthesis is afoot. Rather than producing a wholly original intellectual edifice—a very unusual and difficult achievement in legal thought—the constitutional law scholars have managed to apply some of the most useful insights of legal realism to their theoretical enterprise. In doing so, they have reached some of the same conclusions as realist scholars of contract law and statutory interpretation.

Given this intellectual progress, I believe the moment is ripe for consolidating the gains in the constitutional realm, explicitly seeking parallel truces in the realm of contracts and statutes, and forcing scholars to think more pointedly about which disagreements persist. What appear to be intractable theoretical conflicts are ultimately rooted in differing beliefs about
empirical realities. Such divergent beliefs could potentially be brought into alignment in light of pertinent evidence. In the chapter’s final section, I consider what kinds of inquiries might have the potential to collect such evidence.

The chapter proceeds through its analogy schematically by addressing the following questions for contracts, statutes, and constitutions:

1. What does it mean to draft a textual instrument?
2. What is the process for bringing the text to life?
3. What should interpreters do with unambiguous text?
4. What should interpreters do with ambiguous text?
5. What can formalists and non-formalists agree about? What is the real substance of their disagreements?

Before turning to the analogies, it is worth first considering two disanalogies that might call the present exercise into question. First, contracts, statutes, and constitutions are brought into being through quite distinctive processes: contracts through negotiation between private parties; statutes through legislative action; and constitutions through a fairly elusive and mysterious process, arguably rooted in an act of popular sovereignty. Those bound by a contract have usually affected its features directly and meaningfully, while those people living under the force of a statute or under the auspices of a constitutional system of government usually have not. Indeed, one can make a persuasive case that ordinary people under laws or constitutions have never meaningfully consented, while it is difficult to say the same about most contracts.

I largely avoid this problem by pursuing the analogy only inasmuch as statutes and constitutions govern the behavior of government officials. While thinking about private citizens’ relationship to the law raises many important and difficult questions (what it is to meaningfully

---

3 It is certainly easy to think of many other disanalogies, but for the most part they are untroubling.
4 See Bensel (2007) on the conceptual difficulties of thinking about how any order of rules is brought into being.
consent to be governed, whether there is a *prima facie* moral obligation to obey the law, and other perennial favorites of legal philosophers), the government official’s relationship to the law is generally supposed to be somewhat clearer.\(^5\) Under the rule of law, government officials are to be bound by laws (both statutory and constitutional) in a manner generally analogized to the way a private party is bound by a contract freely entered. If an official openly flouts statutory obligations, they are just as clearly in breach (malfeasant) as a private party who flouts contractual obligations. The fact that the government official need not have been directly a party to the drafting of the textual instrument in question changes little.

Second, there is also the matter of who is considered a party to the textual instrument such that they may seek vindication of their rights within it. Contracts are entered into largely to clarify these questions, with parties self-consciously opting into contract remedies instead of tort remedies. Statutes and constitutions, however, are generally far less explicit about who will have a right of action to enforce the terms of the instrument. Various theorists offer straightforward contract-like answers: the early law and economics literature often contemplated statutes as contracts between a legislature and its interest group patrons (e.g. Stigler 1971, Landes and Posner 1975); later efforts elaborated on this framework, for example suggesting that institutional design choices are made so as to minimize transaction costs between interest groups and the legislature (Horn 1995). Alternatively, others have thought of statutes as contracts between the legislature and the executive branch (e.g. Epstein and O’Halloran 1999), with the president’s veto power giving his branch leverage in the bargaining (Cameron 2000). Neither of these relationships seems genuinely analogous to the one that exists among private parties to a contract, however, because in neither case do they present a set of parties who pursue legal

\(^5\) This is especially so if we stipulate that the state is generally a just and legitimate one. Without saying more about this, I propose to so stipulate for the purposes of this essay, which is motivated almost entirely by thinking about a modern American context in which the rejection of the state’s basic legitimacy is quite uncommon.
claims against each other. In trying to vindicate statutory rights, interest groups do not claim that the legislature has somehow breached their obligations, and only in rare cases do legislators directly pursue legal claims against the executive branch (rather than utilizing political tools of control). Mostly, we see claims made by private citizens (including interest groups) against actions of the executive branch—two parties that never directly bargain with each other. Nor do citizens, as a group, generally possess the power to assert claims under a statute. In the realm of constitutions, there is a hoary debate about who is a party to the social contract—whether it is between a people and their governing officials, or rather between the people themselves. In practice, figuring out who can legally assert constitutional rights and in which contexts turns out to be quite contentious. This difference between contracts (which clearly define the affected parties and the relevant causes of action) and statutes and constitutions (which leave these properties ambiguous) strikes me as a consequential disanalogy, and this issue of standing to assert legal claims under the textual instrument will be consequential as we think through what kinds of practical effects the instruments can be expected to have.6

1. What does it mean to draft a textual instrument?

A. Why do people choose to organize their behavior around textual instruments?

Contract

If two private parties want to interact with each other for their mutual gain, how can a contract help them to do so? Somehow, contracts change the expected value of future actions, such that each party is incentivized to act as the other desires rather than pursuing self-serving opportunism. This allows parties to “project[] exchange into the future” (MacNeil 1980, 4).

6 The discussion in Section 3.B is especially pertinent to these issues.
Like the binding of Ulysses, contractual obligations empower parties to resist siren-songs of destructive selfishness and therefore solve a commitment problem.

But how are the changes in incentives actually brought about? The most obvious and familiar way is to create rights of third-party enforcement: by entering into a contract, party A confers on party B a right to sue (that is, apply for and receive from C a warrant to get restitution from A) under certain circumstances. Reputational incentives may play a similar role. Because transactions are very rarely “one-shot” interactions with no expectation of consequences for future transactions, once a contract is verifiably entered into it incentivizes fidelity to the terms because non-performance will signal to other potential contracting partners that the reneging firm is not trustworthy. This reputational mechanism also provides a form of third-party enforcement, albeit an informal one.\(^7\)

The coordination function of contracts, on the other hand, does not derive its importance from its role in enforcement, but rather in its value for preventing and resolving disputes in the normal course of relations between the contracting parties. By clarifying the partners’ mutual expectations, a contract allows them to make mutual accommodations. In this sense, “law is not what contracts are all about. Contracts are about getting things done in the real world—building things, selling things, cooperating in enterprise…” (MacNeil 1980, 5). These expectations of “effective future interdependence” can become a powerful source of solidarity, especially when stabilized by the promise of third-party enforcement; contract thus allows self-interest to become “one of the most reliable of all bases for solidary beliefs” (MacNeil 1980, 91, 97).

---

\(^7\) Before formal third-party enforcement in state courts was a viable option, the so-called law merchants provided the service of making such reputational signals transparent and reliable (Milgrom, North, and Weingast 1990).
Like contracts, statutes work by changing incentives: most directly, those of government officials. By committing government officials to a standardized and codified course of action, statutes seek to 1) ensure that like cases are treated alike; 2) promote expediency in government action; and 3) make government actions more reliable and transparent, foregoing the “wise, inventive, prophetic” uses of discretion in order to avoid the “outrageous, arbitrary, tyrannical” ones (Llewellyn 2011, 52-4).

How does the commitment mechanism function? Unlike parties to a contract, the officials may not have voluntarily subjected themselves to the law’s authority. This may attenuate reputational incentives created, since violating the terms of a statute might be seen as courageously rebelling against an unwelcome imposition rather than as reneging on a commitment voluntarily undertaken. Even so, government officials are generally at pains to portray themselves as faithful servants of the law, apparently believing that their legitimacy as public servants largely derives from this fidelity.

One might think that a statute’s coordination role would also be limited by the fact that officials were not parties to the law’s creation. And yet statutes do act as coordination devices, creating a shared understanding of how the government will conduct itself. In a sense, statutes are constitutive of government officials’ roles, supplying them with ends to pursue. By publicly

---

8 Naturally, statutes are not the only mechanism capable of mobilizing government. Government officials’ choices could be organized far more loosely, or the “sovereign” could interact directly with citizens through some combination of sticks and carrots. That we now live in a polity in which statutes directing the behavior of government officials are far and away the most common mode of state influence is a reflection of our particular vision of governance rather than a universal truth.

9 This is true inasmuch as government officials are (or see themselves as) agents of the legislature. But, as argued in this dissertation’s Introduction, we should not mistake this for the whole of government officials’ self-conception or real roles.
stating these roles, Congress\textsuperscript{10} clarifies its expectations (at least partially) and the statute’s language provides a reference point for an ongoing dialogue between the branches about the appropriate policy responses to a changing world. Oversight hearings and less formal communications about whether executive branch officials are properly respecting the law, which depend on a mutual desire to come to some workable understanding between the branches and therefore avoid open conflict, often lead to workable compromises.

Finally, statutes also create rights of third-party enforcement.\textsuperscript{11} As discussed above, the question of who has the right to bring claims that government officials have deviated from what is permitted or required under a statute is far less straightforward than in the realm of contract. In some cases, it may be the case that no person or institution is capable of effectively bringing a legal claim to vindicate the law’s demands.\textsuperscript{12} In other cases, however, statutes do create rights of action. If government actions directly harm an individual, that person may have standing to challenge the action as unlawful (either because it perverts the meaning of some law or is unsanctioned by any existing law). Many modern statutes enable citizen suits, in which anyone may pursue legal claims that the government is failing to fulfill its statutory mandate.\textsuperscript{13}

\textsuperscript{10} Throughout this chapter, I often refer to “Congress” as the unitary institution with the sole power to pass or alter statutes, but of course this is an oversimplification. Congress is far from unitary—its two Houses often diverge, minorities in the Senate frequently succeed at obstructing the majority agenda, and the profusion of legislators’ interests makes concerted action profoundly contingent, as the cases in Chapters Four and Five illustrate. Furthermore, as many have pointed out, the United States’ legislative process is basically “tri-cameral” because of the requirement that the president sign legislation. Although formally a Congressional supermajority can override the president’s veto, in practice altering the legislative status quo usually requires the president’s assent. For the purposes of this chapter, however, the benefits of conveniently referring to “Congress” outweigh the costs.

\textsuperscript{11} The term “enforcement” here can be a bit distracting, since we use it most often in reference to the executive branch enforcing the laws against private firms or citizens. In this context, it refers to the ability to enforce the obedience of the executive branch to some understanding of what the law requires. To avoid this confusion, I will often use the word “forcing” to refer to attempts to bring the executive branch into line.

\textsuperscript{12} For example, the Federal Reserve Board of Governors’ responsibility to set monetary policy is governed by a number of statutes, but it is doubtful that any citizen or other institution has the power to take them to court if they believe that the Fed has failed in its mandate.

\textsuperscript{13} See Farhang (2010) for discussion of the related choice of whether to include “private attorney general” clauses allowing citizens to bypass the executive branch in ensuring that a law’s provisions are observed.
Having some kind of a constitution may not be a choice at all—many theorists argue that a “little-c” constitution is inherent in every country’s political arrangements. But having a written constitution is a choice that a polity consciously undertakes. Once again, adopting a textual instrument alters the possibilities for government action, both empowering and constraining government officials.

A constitution does this most importantly through constituting government structures capable of coordinating a society’s politics into organized and reliable forms. Without a basic agreement about these mechanisms for getting things done politically, anarchy and degeneration seem all but certain. A written constitution provides a means by which a society can fix these arrangements purposefully and explicitly, and with the consent of the people through some ratification process, rather than through informal institutional evolution.\(^\text{14}\)

Once the constitution is in place, government officials have very large reputational incentives for acting within its bounds lest they be accused of staging a coup. Because constitutions are unlikely to be overly specific about permissible conduct for government officials, figuring out what actions are truly beyond the pale will rarely be a straightforward matter. A government official’s reputation as a legitimate defender of a constitutional order is therefore likely to be an intensely political matter.

The same can be said about when constitutions will create opportunities for “third-party” enforcement of some constitutional guarantee against a potential transgressor. “Third-party” in this case requires scare quotes because, in some sense, all citizens and governmental actors alike

---

\(^{14}\) Barnett (2004) argues that constitutions serve the same four functions he ascribes to contracts: 1) evidentiary, leaving a definitive record of an agreement; 2) cautionary, promoting reflection for each of the parties; 3) channeling, routing disagreements into well-established forms; and 4) clarifying, requiring precise articulation of what is agreed upon. Morrison (2005) powerfully critiques this position as foreclosing the possibility of open-ended constitutional provisions and dynamic interpretation merely on the basis of assertion.
are constitutional actors, with no real recourse to genuinely disinterested third-party adjudicators likely to be available. Instead, a government must enforce constitutional rights against itself—a situation which has inspired commentators all the way back to Plato to ponder who can guard the guardians. In an American separation-of-powers style system, the knot is cut by pluralizing the government, so that one group of institutional actors can stand as a bulwark against the others’ alleged constitutional transgressions. If the judiciary is seen as a genuinely coequal constitutional actor, then there can be little doubt that judges must refuse to sustain a legislative or presidential action if it directly contravenes the supreme law instantiated in the constitution, leading to the process we now know as judicial review.\textsuperscript{15} Once again, because there is no strictly neutral party available to enforce the one true meaning of the constitution, questions about which actions are actual violations are likely to have an intensely political element, but this need not trigger degeneration into chaos.\textsuperscript{16} As with statutes, the question of who will have the ability to successfully assert claims about constitutional requirements will be a complicated one. Even having registered all these caveats, creating certain enforceable limits remains one of the most powerful motivations for a people to bind itself to a written constitution in the first place.\textsuperscript{17}

**B. How much can the textual instruments resolve? What is left unsettled?**

Having outlined what people hope to gain through committing themselves to some textual instrument in the first place, we must now ask what purposes those instruments are capable of constructively and effectively serving once they are commenced. In each case, the answer is that the textual instrument can effectively decide a great deal—but only in limited (and

\textsuperscript{15} Hamburger (2008) shows quite persuasively that the way judges at the founding (and before) thought of this responsibility does not cleanly match up with our modern idea of judicial review, but for our purposes here the basic idea of judges owing their ultimate fealty to higher law remains stable.

\textsuperscript{16} See Paulsen (2003, 2378-79).

\textsuperscript{17} See Whittington (1999a, chap. 5).
generally less important) circumstances should we expect it to be capable of providing a resolution to every dispute. We should also be wary of official pronouncements about how much the rules in textual instruments are capable of deciding. Because, as advocates, lawyers must pay homage even to sloppy laws, poorly constructed rules end up getting credit for being far more “guidesome” than they really are (Llewellyn 2011, 47). Although it is not the most analytically thrilling maneuver, when assessing the question of how much textual instruments can resolve, the answer must be: “a lot, but far from everything.”

Contract

A contract specifies who owes what to whom, when. If the parties to the contract were in a reliable and static environment, such a specification might well be sufficient to prevent any disputes from arising, or to resolve them clearly if they did arise. The world, however, is rarely reliable or static. As a result, parties are unlikely to explicitly account for every possible contingency in their written contract. They may still formally make their contract complete by agreeing to default rules which specify appropriate responses to situations to which the text does not directly speak. Nevertheless, in reality the accession to default rules is far from total, leaving a large portion of contracts significantly incomplete because of the transaction costs of achieving full agreement combined with parties’ bounded rationality (Simon 1951; Williamson 1985). Adding more and more clauses to the contract that determine mutually acceptable responses to contingencies is prohibitively costly, and so contracts are left incomplete.

As a result, it is generally a mistake to understand contractual partners’ entire relationship in terms of the written instrument. The closer a transaction is to being purely “discrete”—that is,

18 As MacNeil (1980, 8-9) describes it, achieving a complete contract is generally made impossible by the limited nature of our minds, our highly imperfect knowledge of the future, and “the need to fit them into symbolic forms required for communication.”
a one-time exchange which is commenced and completed in a short period of time—the more likely it is that the written contract will be able to include all the parties’ concerns. As relationships become less discrete—that is, drawn out over time, repeated, and susceptible to more unexpected changes in conditions—written contracts are likely to become less accurate representation of the real governing norms of the relationship (MacNeil 1980). Just as there can be unwritten “little-c” constitutions, so too are there unwritten yet real aspects of a contractual relationship that elude being neatly captured in words. Reputation, coordination, and even third-party enforcement are likely to go beyond the written terms, and the “real” contract—rather than the “paper” one—can be thought of as “whatever the officials do about promises…” (Llewellyn 1931, 717). This may include aspects of the relationship unrecorded in the textual instrument; it may also fail to include some of what is a part of the contract.

Statute

Like contracts, statutes set out who must do what.¹⁹ In principle, a legislature might be able to write statutes completely specifying every aspect of the required actions, including indication of exactly which government officials are responsible for ensuring that those actions are realized. In practice, legislatures rarely come anywhere near completeness, for they are unlikely to be either interested in or capable of doing so. The transaction costs of acquiring the information necessary to a complete statement of rules are usually insuperable, especially if the requirements rely on technical standards or intricate knowledge of policy details. More fundamentally, though, individual legislators receive paltry rewards for mastering the knowledge

---

¹⁹ Clearly some symbolic legislation resists this description. Admittedly, my frame of reference centers on regulatory statutes.
necessary to write more detailed and complete legislation. At the same time, legislators collectively have little incentive to expend energy learning about, arguing over, and finally compromising on every aspect of their rules. What Sunstein (1995) calls the “incompletely theorized agreement” is the lifeblood of modern government, without which it is doubtful anything could ever be accomplished. As they set down instructions for executive branch actors in their statutes, legislators are not required to reconcile their possibly-conflicting ends in any determinate way. It can leave this work to the bureaucrats or judges, and in many cases doing so is extremely politically convenient.

Sometimes, Congress might clearly delegate some choices to particular actors while explicitly requiring that their choices about certain parameters remain within a prescribed zone. The political science literature about the relationship between an enacting legislature and bureaucrats charged with administering laws generally treats this arrangement as prototypical (at least when “delegation” is involved) (e.g. Huber and Shipan 2002). In this vision, the Congress chooses the tens and ones places of a policy, but allows bureaucrats to fill in the decimal places however they see fit. The statute is incomplete, but only in its minor premises. Few statutes fit

---

20 As Arnold (1990) explains, effective policy is only electorally valuable for legislators to the extent that it might change the world in ways that voters would notice, care about, and connect to the legislator. Compared to the electoral reward of direct engagement with constituents, this is quite attenuated. Of course, there may be legislators who find legislative craftsmanship rewarding for reasons beyond electoral rewards. In Chapter Five, I argue that the attention of such “legislative superintendents” is likely to be crucial for the creation and maintenance of effective policy.

21 Throughout this dissertation, I use the word “indeterminate” to mean “not fully determinate,” although some scholars have argued that this meaning is better captured by “underdeterminate,” with “indeterminate” reserved to mean totally devoid of objective meaning—see Solum (1987). My choice is largely an aesthetic one; what Solum calls “indeterminate” I call “completely indeterminate.”

22 The passage of the Sherman Act provides an excellent example. In the late 1880s, the public had become widely incensed at the rapid rise and perceived abuses of the “trusts,” and all agreed that something should be done. “Monopoly” was an evil that all could rail against, and “competition” a desirable value to embrace—regardless of whether these terms had any well-specified meaning. As Letwin (1965) persuasively argued, members in Congress had little motivation or ability to decisively determine what the proper balance was between allowing free transactions and restricting them in order to promote unfettered competition and in the end abstained from trying to do so in any kind of specific way. Instead, they concentrated the law’s power in almost absurdly broad provisions prohibiting the “restraint of trade” and “monopolization” and left it to the executive and courts to determine what these phrases would mean in practice.
neatly into this model, however. Instead, most statutes instantiate requirements through verbal formulae that cannot be given any practical meaning without some act of interpretation. (This is the case for nearly all laws, and not merely those containing any self-conscious “delegation.”)

Since statutes are not self-interpreting, we should not expect statutes to comprehensively answer all of the questions about the policies they require or initiate.

**Constitution**

Finally, how much do constitutions decide about the operation of the political systems they send into motion? Once again, it can specify what is required of whom—but in this case the aspiration will never be to provide a complete ordering. Doing so would require extinguishing future politics, rather than enabling constructive political engagement.

If all constitutions leave spaces, however, some might argue that a constitution is nevertheless capable of specifying with some exactitude what those spaces consist of, with each government actor given a precisely circumscribed domain. In principle this seems right—and many constitutions of recent vintage seem to take this idea to heart, providing dozens or hundreds of pages of specific requirements and limitations. The American constitution, however, clearly takes a different path. Congress’s powers are limited to enumerated categories, but few of these have a self-evident scope and the listing importantly includes the power “to make all laws which shall be necessary and proper for carrying into execution” the others. The framers quite consciously debated the merits of including such an open-ended warrant, ultimately deciding that the flexibility it conferred would be vital to a government capable of adapting to a changing world.\(^{23}\)

---

\(^{23}\) Beeman (2009, locations 5229-5234, 5495-5512). See also *McCulloch v. Maryland* (1819, at 407), in which Chief Justice Marshall noted that the framers consciously avoided limiting federal powers to those “expressly”
2. What is the process of bringing the text to life?

A. Experience

If text does not straightforwardly resolve every question, what are the processes through which contracts, statutes, and constitutions acquire their practical meaning? Holmes’s (1881) famous adage is a good start: “The life of the law has not been logic; it has been experience.” Textual instruments come to mean what the people who use and rely on them believe they mean.

Contract

I can do little to improve on Karl Llewellyn’s description of how contractual terms are animated through practice. A contract’s written terms provide the parties an original framework, a constitution, a source of ultimate sanction in dispute or break-down…

But for the running of affairs they say but little. The play of personality, the unrecorded adjustment from day to day, further factual agreement from time to time, informed by usage, and by initiative and acquiescence which do not even call for conscious agreeing—these are what fill the contract frame-work with a living content; these are what often so stretch and overlay it as to make the initial contract a wholly misleading guide to what occurs. (Llewellyn 1931, 730-731)

Generalizing about all rules, Llewellyn believed rules must be understood as something more than the words on a page; “the propositional form of a rule is one thing, what the rule actually does (and, thus, is) is another” (Schauer 2011, 7-8).

Contractual relations thus come to be richer and thicker than the textual document that serves as the relationship’s anchor. Norms beyond those specified in the text, such as role integrity, contractual solidarity, and flexibility come to inform the parties’ behavior toward each other, so that “[a]s contractual relations expand, those relations take on more and more the

---

named because of the experience under the Articles of Confederation, and explained the broad nature of the Constitution thusly: “A Constitution, to contain an accurate detail of all the subdivisions of which its great powers will admit, and of all the means by which they may be carried into execution, would partake of the prolixity of a legal code, and could scarcely be embraced by the human mind. It would probably never be understood by the public. Its nature, therefore, requires that only its great outlines should be marked, its important objects designated, and the minor ingredients which compose those objects be deduced from the nature of the objects themselves.”
characteristics of minisocieties and ministates,” and the external norms of the broader society are often internalized to the relation (MacNeil 1980, 70; Kornhauser and MacLeod 2010).

Although the expectation of performance of the contractual terms obviously underlies the parties’ willingness to contract in the first place, real performance nevertheless subtly alters and fills gaps in the textual requirements. “Even meticulous performance of the most explicit planning transforms figments of the imagination, however precise, into a new, and therefore different, reality. A set of blueprints and specifications, however detailed, and a newly built house simply are not the same” (MacNeil 1978, 873). 24 Of course, nonperformance also alters contractual relations—no matter how empowered the other party is to respond—by shifting expectations about future conduct (MacNeil 1978, 874).

Any of these changes could, in theory, be captured through a formal agreement to a re-drafted instrument, but in practice, this formality is often dispensed with. Courts are likely to give some weight to patterns of actual conduct and reliance when these are at odds with the contract’s explicit terms, suggesting that even in a legal sense the contract’s real meaning is thought to derive from something more than what is contained in the text.

Statute

The reality of a statute also depends upon the ways that it is put into action, in this case by government officials. Many perfectly-operative-sounding laws do little more than gather dust in the law books, failing to serve as actual guides for action. 25 On the other hand, laws that

24 MacNeil (1978, 894) also explicitly drew the analogy between contract and constitution: “I feel some temptation to think of the written parts of contractual relations, especially very formal parts, such as collective bargaining agreements and corporate charters and bylaws, as constitutions establishing the legislative and administrative process for the relation.”

25 Of course, as the discussion in Chapter Four makes clear, this does not foreclose the possibility that some enterprising government official may later rediscover the law and figure out how to liven it up for his own purposes.
inarguably organize government behavior do not always do so in the ways that their text might seem to prescribe. Certain parts of the statute may receive more attention than others, and daily administration of the law may turn the law to purposes unanticipated by its authors. Indeed, in many cases a sudden shift to following the letter of the law as it is on paper might well paralyze the actual operation of government; “going by the book” is a stock phrase used to describe a frame of mind that is not altogether natural (see Bardach and Kagan 1982).

In recognition of the fact that working statutory regimes consist of more than just statutory text, some scholars have suggested thinking about “policies as institutions” (Pierson 2003). Usually there is a statute at the heart of these, again without the suggestion that the written statute tells the whole story. Following the usage of contemporary constitutional law scholars (discussed presently), it makes sense to speak of statutes as providing a framework for government action, the substance of which will put flesh on the skeletal structure provided by the statute. This perspective gives a much more prominent place to the executive branch in thinking about what statutes really mean, since it is their quotidian development of details that determines what will be made of a statute (thus making their contribution in toto rather more than just the filling in of a few incidental details). We are likely to see what are basically expectations-based common laws evolve around executive branch actors’ decisions (see Sunstein 1998).

**Constitution**

Contemporary scholars of constitutional law have increasingly emphasized the limits of simply following the text as a way of determining constitutional practice. While they emphasize that the text must be capable of exerting some constraining force—because only some

---

26 Recent American Political Development literature has focused on the ways in which policy may be “retrenched” even without legislative changes; see especially Pierson (1994) and Hacker (2004), and on a somewhat different note Patashnik (2008).
constitutional interpretations are actually defensible—they also acknowledge that interpretation of the text cannot answer many questions. Beyond these limits of interpretation, constitutional construction—a related but distinct practice, in some sense requiring reliance on outside (and contestable) political commitments—determines the features of the political system that the document constitutes (Whittington 1999a, 1999b).

With its characteristic brevity and reliance on open-textured phrases, the American Constitution operates in large part as a basis for indeterminate political construction. The document serves as a coordination mechanism for organizing the development of politics rather than as a source of determinate legal interpretations. This means that while the document does not decide issues specifically, it nevertheless sets the terms of debate and discussion about the shape and character of the national government. Allowing the Congress to impeach the president for “high crimes and misdemeanors,” for example, provides guidance but no sure answers about what kinds of transgressions are impeachable. As a result, working out the meaning of this language was a function of political struggle as much as the text (Whittington 1999b). Balkin (2009) promotes the metaphor of constitutional “frameworks,” rather than fully built “skyscrapers.”

B. Resolution of Conflicts

Those working with a textual instrument will do much to work out its practical meaning through use and through incremental adjustments, but this is not to say that all can be expected to go smoothly. When mutual accommodation proves impossible, the third party adjudicator takes the leading role in determining just what the contract, statute, or constitution really requires.

27 Balkin apparently arrived at his architectural metaphor without any consciousness of its usage by contract scholars, but it is a telling parallel.
Especially in American society, legal meaning is often developed through a process of “adversarial legalism” (Kagan 2001).

Contract

Once the parties are unable to resolve their own disagreements, they give themselves over to the law for an authoritative resolution. Note that if it is completely clear what the contract requires, and thus how a judge would rule, then the party who is sure to lose has little to gain by contesting the point. Consequently, most disputes that reach a court will involve claims with at least some merit on both sides and some uncertainty in how a judge is likely to rule. Again, the words of the contract prove incapable of exerting a real effect in and of themselves. Llewellyn pithily explains: “A contract is no equivalent of performance; rights are a poor substitute for goods” (Llewellyn 1931, 724). A court’s ruling resolves the present dispute while also setting a precedent for future ones, so that if the parties to the contract are able to maintain their relationship in spite of their legal fracas their contract will be more certainly defined as they move forward.

Statute

Like contracts, statutes are not self-interpreting, and often take their final shape in the crucible of legal contestation. In Llewellyn’s distinctive idiom:

‘[T]he law’ is not self-operative, in the pinch, in regard to the ‘legal consequence’ the rule lays down; that ‘legal consequence’ is an officially stated ‘should be’, but in the pinch it can be translated into ‘isness’ only through a human being in office. (Llewellyn 2011, 39)

28 Persevering in their claim does, however, allow them to impose the costs of an extended legal process on the other party to the contract, which they may use as a source of leverage in negotiating a settlement more favorable than the result a court would likely impose.
In other words, where the meaning of a statute is legally disputed, it is the human being in judicial robes who ends up resolving just what that meaning is.

Chapter Three’s case study of the development of the Clean Air Act shows this dynamic at work. There, interested parties dissatisfied with the EPA’s reading of the statute indicated their displeasure through adverse comments. Often these led to some kind of voluntary adjustment by the agency; just as often, they were affably shrugged off with some indication of why the agency believed their interpretation was defensible (or even correct). In many instances, unresolved disputes led to the initiation of litigation, and ultimately the D.C. Circuit became responsible for deciding just what the statute would mean in practice.29

As judges resolve these disputes over statutory interpretation, should we think of them as anything other than surrogates for the political parties of the presidents who appointed them? As Chapter Three discusses, the existing literature as well as my own data certainly suggest that there is a role for standard partisan differences, but to my eye they do not suggest that we can call this a primary determinant.30 Even if judges are not merely “legislators in robes,” as some would label them (e.g. Segal and Westerland 2011), we may nevertheless find it somewhat unsettling to find that statutory meaning “in the pinch” is a function of how particular cases—with particular attorneys and particular judges—are resolved. “The law,” often conceived of as something fixed and determinate, is at least partially an outcome of a dynamic process which involves advocates, sometimes of unequal skill, doing their utmost to serve their client’s interests by manipulating the “machinery of the rules” within the “leeway” of respectable limits, and judges who, while guided primarily by the law, also take into account the “health (as he feels it)

29 For another study finding that litigation is a regularized part of the process giving practical meaning to statutory text, see Coglianese (1996).
30 Tamanaha (2009, Ch. 8) offers a more comprehensive review of empirical studies of political judging in appellate courts and reaches the same conclusion. Also see Edwards and Livermore (2009, 1910-18).
of our legal scheme of things and justice (as he feels it) in outcome of the adjudication in hand” (Llewellyn 2011, 119).

**Constitution**

Finally, as indicated in Section 2.A, constitutions by their nature leave much about their practical meaning to be determined through contestation. Unlike the situation for contracts and statutes, the identity of the ultimate arbiter of constitutional disputes is itself a source of contestation. The modern folk view is that the Supreme Court simply has the last word when it comes to deciding issues of constitutional meaning; in the Court’s own most grandiose pronouncement, of *Cooper v. Aaron* (1958), it sets itself up as possessing the single supreme power “to say what the law is,” regardless of what any other government officials (or the general public) might have to say about constitutional meaning. As Whittington (2007) shows, this formulation of judicial supremacy has often suited the needs of presidents and legislators, but their devotion to the idea is unlikely to survive if it becomes a major hindrance to their political agenda. Major conflicts about constitutional structure have incontestably led to huge changes, cheered on as legitimate exercises of popular sovereignty and dubbed “revolutions” by their defenders such as Ackerman (1993) and denounced as illegitimate departures from the constitutional text by critics.\(^\text{31}\)

Having said this, the Supreme Court does often have the last word in settling disputes about constitutional meaning, and many of the concerns about contingency in the statutory realm apply with equal force to the constitutional context as well. Irons (1982) persuasively argues that the key cases of judicial review during the 1930s were decided largely on the basis of the performance of their legal counsel. The outcome in *Brown v. Board of Education* (1954), the

\(^{31}\) For a penetrating critique of the constitutional structure forged during the New Deal, see Greve (2012).
twentieth century’s most celebrated Supreme Court ruling and a unanimous decision, seems to have depended largely on the legal strategy carefully developed by the NAACP over the decades leading up to it (Kluger 1976; Tushnet 1987). Epp (1998) and Teles (2008) both emphasize that there must be a “support structure” of capable lawyers to consolidate any rights that courts might declare constitutionally protected, and each highlights the political economy of growing and sustaining such networks. Even in the weightiest constitutional matters, then, the particulars of legal contestation—what are the cases, who are the lawyers—will matter in determining the effective meaning of the constitutional text.

3. What should interpreters do with unambiguous text?

As Section 1.B discussed, if textual instruments are to provide reliability, they must have the potential to bind parties in deed as well as in word. For this to be true, if interpreters confront genuinely unambiguous language they must do what the text demands (and avoid doing what it prohibits). It seems that the rule must be: “Just follow it!” In each of the areas, this principle earns enthusiastic endorsement—but nevertheless captures only a part of the more complex reality, in which obedience to the text competes with other values and can only be enforced under certain conditions.

A. Endorsements of “Just Follow It!”

Constitution

I take the three areas in reverse order here because it is in the constitutional realm that we find the American locus classicus of endorsement for following text: Marbury v. Madison (1803). Although it is quite misleading to imagine that Chief Justice Marshall somehow
“invented” judicial review in this case, his statement of the obligation to obey the supreme law nevertheless stands as epochal. As Paulsen (2003) has pointed out, Marshall’s opinion need not be read (as it is by most contemporary readers) as arguing for the supremacy of the judiciary’s understanding of the constitutional text. Rather, the argument is for “constitutional supremacy—the supremacy of the document itself over misapplications of its dictates by any and all subordinate agencies created by it” (2709). Where the constitution is perfectly clear, fidelity to our system of government requires that its dictates be followed. To dispute this proposition (rather than particular points about which language really is clear, or what clear language requires) is, in essence, a rejection of the idea of a written constitution (see also Whittington 1999a, chap. 3).\textsuperscript{32} Where actors in various institutional contexts embrace the “Just Follow It” attitude \textit{and} they can all agree about what the constitutional text clearly requires, obedience to the text should prevent conflicts from arising.

\textit{Statute}

When we turn to the statutory context, the logic is nearly identical. Unless a statute is to be mere window-dressing, it must be capable of binding government actors. The first “step” of the test famously laid out in \textit{Chevron v. NRDC} (1984)\textsuperscript{33} echoes Marbury’s endorsement of textual supremacy: when reviewing an agency’s interpretation of a statute it administers, the court must begin by asking “whether Congress has directly spoken to the precise question at issue. If the intent of Congress is clear, that is the end of the matter; for the court, as well as the

\textsuperscript{32}There are some scholars who are willing to walk this path, but they are likely present themselves and be received as “radicals,” probably for good reason. For example, see Unger (1986), for an extended development of the idea that law is a struggle for power and those who are disadvantaged should do whatever they can to get the most from the system, all fictions of proper legal etiquette aside.

\textsuperscript{33}Throughout the dissertation, full citations to judicial opinions are listed chronologically after the main text of the chapter.
agency, must give effect to the unambiguously expressed intent of Congress” (842-3). To the extent that the law is truly binding, the judiciary must be ready to disallow actions that are not in accord with their understanding of its meaning.\textsuperscript{34} Anyone claiming in the abstract that statutes’ clear meaning should not be controlling seems to be rejecting something fundamental about our system of government. Where executive branch actors understand the law’s clear instructions in the same way as potential legal challengers and judges, no legal disputes should arise.

*Contract*

Finally, the situation in contract law is much the same. Where the text of a contract clearly requires performance of some obligation, if a court is asked to adjudicate a dispute it must enforce that obligation. Given that cooperation between the parties has broken down, and they find themselves as adversaries in court, “the legal aspects of the bargain take on peculiar importance” (Llewellyn 1931, 731). Courts’ privileging of contractual text is encapsulated in two overlapping rules of contract law, the parol evidence rule and the plain meaning rule.\textsuperscript{35} Together, these stand for the proposition that: “A court will refuse to use evidence of the parties’ prior negotiations in order to interpret a written contract unless the writing is 1) incomplete, 2) ambiguous, or 3) the product of fraud, mistake, or a similar bargaining defect” (Posner 1998, 534). Once again, there are likely to be difficult questions about what makes a contract

\textsuperscript{34} As in the constitutional case, subordination to the commands of a clear textual instrument is not really equivalent to subordination to judicial interpretations of that text. Of course, if clear is really clear, then this distinction should not matter. In practice, the executive branch rarely commits itself to interpret some arguably “clear” text in direct opposition to a federal court’s interpretation—but there have been instances, and the practice even acquired a semi-respectable status in the late 1980s under the august-sounding name, “non-acquiescence” (Estreicher and Revesz 1989).

\textsuperscript{35} Barnett (2004, 102-3) argues that where the Constitution should also be interpreted with a strong parol evidence rule whenever its text can be understood as a complete, while supplementary evidence must be admitted where the document’s text is clearly incomplete.
complete, clear, and procedurally sound, but few will doubt that clear language, when available, should be controlling.  

B. Limitations

Contract

A third-party adjudicator of a contractual dispute will follow clear contractual language, then—but this is far from saying that clear contractual language will always be given effect. When deciding whether to pursue formal enforcement, a party to a contract must weigh what can be gained through a vindication of legal rights against what will be lost by turning to the courts, both in terms of the costs of making one’s case (in time, energy, and legal expenses) and in terms of potentially destroying goodwill with the party being sued. If the counterparty can be easily replaced by a competitor, this last cost will be minimal. But Williamson (1985) explains how economic relationships beginning as anonymous market interactions often come to resemble bilateral monopolies as the parties make relationship-specific investments. In a famous example of this “asset-specificity” provided by Joskow (1985), someone looking to build a profitable coal power-plant would ideally like to site it very near a coal mine to minimize transport costs. Making this choice unfortunately puts the new power-plant owner at the mercy of the coal mine, which can refuse to sell at any price lower than the nearest competitor could offer, complete with transport costs. Writing a contract that can be enforced by a court provides only a partial solution to this problem, because behavior will still be what Williamson calls “opportunistic”—basically, the parties cannot depend on fidelity to the contract if there is no viable and cost-

---

36 As Chapter Three points out in the statutory context, given an expectation that courts will indeed enforce clear text, few lawsuits contesting the application of clear language should ever proceed; cases argued before judges will be a completely unrepresentative sample of working law. See Priest and Klein (1984) for further thoughts about case selection in litigation.
efficient way of punishing defection. Enforcement of contractual terms, no matter how clear, is always a costly affair, and it is a mistake to imagine that those with contractual rights will always prefer to pay to see them vindicated.

Statute

Giving a statute’s requirements real legal bite requires more than catching government officials (apparently) flouting its requirements. A party with such a belief must also have standing to bring forward a legal claim. Except for extremely special cases (e.g. McConnell v. FEC (2003)), the legislators responsible for writing the statute cannot themselves initiate court involvement. Instead, in a situation not analogous to the private context, some other party with standing must do so (as discussed in the introduction).

Assuming someone can (and does) bring the matter in front of a federal judge, the next question to ask is what sort of demand the law makes that the executive is allegedly flouting. Depending on the issue, executive obedience to such a demand may or may not be justiciable. Even if the judge is in a position to find an agency’s action (or inaction) unlawful, she may not possess effective legal sanctions to compel the executive to alter behavior. Injunctions may be issued (or in extreme cases criminal penalties for malfeasance applied), but at the end of the day the agency may be able to pursue the policy in question with only superficial changes. If judicial power is lacking in a particular case, no one may find it worth pursuing the protracted process of litigation to begin with. Alternatively, the party who disapproves of the way a government

---

37 They can, however, initiate quasi-legal contempt of Congress proceedings, though they have rarely pursued this option in recent years. See Chafetz (2009).
38 From a purely formal perspective, not acting where a statute specifically demands action is an example of behaving contrary to the statute, and so one can argue that non-performance is simply a species of contrary performance. Legally speaking, however, there are different responses to non-actions than to clearly contrary actions (Biber 2008), and so the two situations deserve separate analysis.
agency is interpreting what it believes is clear language may try to advance its position through
direct engagement with the agency or through political pressure independent of any legal
claims.39

Constitution

No matter how clearly it is stated, a constitutional provision needs someone willing and
able to insist that it be followed if it is to be effective. The situation is much the same as with
statutory requirements, although constitutional standing follows different rules. In most cases,
only those citizens directly harmed by some government action will be able to bring a
constitutional challenge in the courts (usually claims that some protected right has been
improperly violated), creating a large class of government activity which is immune from legal
challenge.

Where legal challenges to enforce a constitutional provision are impossible, the
constitution’s famous “checks and balances” enable political enforcement of clear constitutional
language. If opposing ideas about what is “clearly” required or permissible under the
constitution prove irreconcilable, there may be no averting a sundering of the polity or a violent
conflict in which one side’s interpretation prevails by virtue of military might.40

39 On the other hand, engaging in litigation where one’s claim seems sure to lose is not necessarily irrational. While
litigation is often pursued in hope of direct results, it is often pursued for political spectacle, and in some
circumstances a loss in court could conceivably be as valuable as a victory by changing the politics of an issue.
40 Clearly the American Civil War fits this description. Graber (2006) attempts to explain how reconciliation and
the preservation of constitutional “peace” were live possibilities and that the now-vilified Dred Scott (1857) decision
was actually a valuable step toward realizing them—but this is a controversial position, to say the least, and one
wonders if Graber underestimates the intensity of feeling separating the combatants.
C. Exceptions

Contract

For all the apparent consensus that clear language must be followed, clear language is not, in fact, always followed. Each respective domain has a number of doctrines allowing exceptions that seek to enable courts’ to go beyond the letter of the law in accommodating the complexities of the world. Sometimes this process proceeds somewhat covertly. Confronted with what they consider to be a manipulative contract, judges may engage in a practice of “‘construing’ the particular language in question not to have intended the result it did intend” (Llewellyn 1931, 732). This allows judges to pay homage to the principle that clear contractual agreements will be honored while simultaneously failing to honor them, instead helping the disadvantaged party.

Alternatively, courts may explicitly declare that they are setting aside explicit language for a number of reasons. One is that, as described in Section 2A, a contractual relationship may evolve over time such that the “real” terms of interaction no longer resemble the paper ones. A party may come to rely on ex post adjustments, even where clear contractual language seems to require something else, if the changes were mutually beneficial, actually acceded to by both parties, fair, responsive to some unanticipated events (MacNeil 1978). “Neoclassical” contract law takes these factors into account through doctrines of consideration, impossibility of performance, frustration, and mistake (all of which look to “pick up the pieces” of a terminated relationship) and good faith (which attempts to facilitate continued reliance) (MacNeil 1978, 875, 878-80, 899-900).

Traditionally, chancery courts devoted to equity rather than law were able to enter equitable injunctions against judgments under the common law in the interests of justice. This
meant that a judgment enforcing a contractual requirement could be set aside under certain circumstances (Gitelman 1995, 225). While a few jurisdictions (importantly including Delaware) still have separate courts for such actions, most American state courts of equity were merged with common law courts of general jurisdiction early in the country’s history, and Congress chose never to create separate federal courts of equity (McCormick 1928). Modern American courts can thus rule either on the basis of law or equity and can offer equitable relief in civil cases in the form of injunctions, specific performance, or other non-monetary form, showing how courts seek to uphold values beyond simple obedience to the text of formal legal documents.\textsuperscript{41}

\textit{Statute}

In a number of exceptional situations, courts are also inclined to disregard statutory language, notwithstanding its apparent clarity. As with contracts, judges can accomplish this surreptitiously by simply reading the statutory language in a creative, unintended way while insisting that they are merely “following” the law’s clear requirements.\textsuperscript{42}

In other cases, judges will forthrightly set aside statutory language. Under the doctrine of scrivener’s error, they will do so if they believe there has been a mistake in the process of drafting so that the law as codified diverges from the law that the legislature intended to enact. Few judges adhere to such a strict version of textualism that they will insist on treating codified language as “the law” even when convincing evidence can be marshaled to show that the

\textsuperscript{41} Emmerglick (1945) reflects on the tensions inherent in law and equity, arguing that there is a deep societal need for some discretionary mechanism capable of putting aside results reached through strict observance of law. \textsuperscript{42} Alexander and Sherwin (1994) argue that all rules are conceived with the understanding that they will be disregarded in cases where following them will lead to egregious injustices, in spite of authority figures’ frequent insistence (the authors advisedly call it a “lie”) that rules are “serious” and to be treated by everyone as authoritative all the time.
language was a result of poor draftsmanship that frustrated clear intent (as opposed to a sneaky or indeterminate formulation that some faction included to make compromise acceptable).

Rather than insisting that legislative errors are Congress’s to repair, most judges will simply approach the statute as if it had the supposed cleanly drafted form. Justice Scalia, for example, has argued that a literal meaning should be rejected if it “obviously contradicts the statutory intent” (cited in Siegel 2001, 331).

Quite similarly, even fairly hard line textualists often admit exceptions for absurd results. Rather than assuming that Congress failed to successfully express its intent, the absurd results doctrine says that even Congress’s clear expressions should not be allowed to compel an absurd result. What exactly is so outrageous as to be “manifestly” absurd is a question difficult to answer directly, but the concept is often defined elliptically through canonical results. In general, it seems the doctrine is used to protect values of reasonableness, rationality, and common sense (Dougherty 1994).

Rather than protecting those values through arguments about absurdity, judges could alternatively act directly on their behalf. Citing the English common law doctrine of the equity of the statute, some scholars argue that judicial responsibilities properly include acting on behalf of equitable concerns even when their decisions contravene clear statutory language (Manning 2001). Under this conception, judges are something more than faithful agents of the legislature. This might be problematic for the rule of law in the American context of separated powers, as

43 A notable exception is Manning (2003), who urges textualists to have faith in the systemic value of their creed even to the point of apparently absurd results in individual cases. Manning believes that if textualists are context-sensitive and sophisticated in their reading of statutes, rather than stubbornly literal-minded, this will produce few really troubling results. Siegel (2001, 327-29) provides a strong case that any reliance on an absurd results doctrine opens textualism to intentionalism, allegedly the textualist’s bane.

44 There are two especially frequently cited examples: first, Pufendorf’s citation of a Bolognian statute prohibiting the spilling of blood on the streets, which was not applied to a surgeon who opened the vein of a person having a fit; and second, a statute of 1st Edward II, making an escape from prison a felony, which was not applied to prisoners escaping a burning jail, “for he is not to be hanged because he would not stay to be burnt” (Dougherty 1994, 139).
Manning argues, but then again it might be taken to promote values of coherence and fairness that are also integral components of the rule of law (see Dougherty 1994, p. 131). Similarly, under the Administrative Procedures Act, judges are asked to set aside administrative actions that are “arbitrary and capricious.” If an agency can really show that it is “just following the law,” then presumably a judge would not find its actions arbitrary, but given a result in which an agency seemed to be wantonly trampling a citizen’s interests the judge might instead find that the result did not really follow from the statutory language, at least without better justification.

Constitution

Finally, clear constitutional language is not always likely to be taken at face value, either. Justice Hugo Black (1960, 874) famously asserted that when the first amendment says that “Congress shall make no law…,” that this really means no law in the relevant areas, but this is a decidedly minority view. Given that the government can show a compelling need, the categorical language of rights is routinely set aside. Constitutional language is treated as constitutive rather than absolute and inviolable, which is not to say that any old concern can displace clear constitutional requirements. Inferior court judges ruling on issues of constitutional law also seem quite reluctant to set aside Supreme Court precedents even when they believe those precedents to be inconsistent with the plain meaning of the constitutional text, apparently demonstrating that they are privileging other institutional values over fidelity (Levinson 1995, 505-6).

---

45 Fried (2000) compares the use of strict scrutiny in the constitutional context to the doctrine of scrivener’s error in the statutory context.
46 Whittington takes as one of his premises “that the purpose of judicial review is to interpret the Constitution”—recognizing that other purposes could be offered but noting that scholars are generally reluctant to frame their arguments in non-interpretive terms (1999a, 2). As this section and the conclusion make clear, however, I find this to be more of a best-case scenario rather than a universally defensible presumption.
All of these exceptions recommend suspicion toward strong metaphysical arguments which say that “the law” and “the text” are one and the same. Since text—even clear text—acquires force only through practice, it seems that when push comes to shove few students of the law insist that we feel collectively tyrannized by the text in cases where we all feel that it is serving no deeper social purpose.47

4. What should interpreters do with ambiguous text?

A. Who decides?

Contract

When parties to a contract are confronted by ambiguous text, they each have their own interests in determining what is made of this text. Each can be expected to seek her own advantage, although presumably the contract is meant to serve their mutual advantage. In many cases, faced with the uncertain prospect of taking disputes about ambiguous language to a third-party adjudicator, the parties will manage to come to some mutually acceptable understanding about the meaning to be given to the ambiguous text. As described in Section 2, the practical meaning of the text will come to be defined through practice and the informal resolution of conflicts. When conciliation proves impossible, the parties will be forced to treat the third-party adjudicator’s decision about the meaning of the text as authoritative—or at least take it as the new baseline for informal mutual accommodation.

If the parties are likely to have a difficult time accommodating each other whenever disagreements about ambiguous language arise (i.e., if the ambiguous language is likely to make

47 Of course, this statement merely reframes the problem rather than solving it. In the euphemism of “we all” comes the rub, since unanimity is a rare commodity where interests are potentially in conflict. Where some feel the text of the law is tyrannical but others do not, there are good reasons to let clear text be definitive, discussed in Sections 1A and 3A.
the transaction costs of the ongoing relationship unacceptably high), a different strategy is to specify ex ante which of the parties will have the authority to resolve the meaning. Such “residual decision rights” essentially elevate one party to the contract, leaving the other party with only those legal rights explicitly and clearly specified in the contract. Owners typically use such contracts, giving lessees certain enumerated rights to use their property but reserving power to decide the residual questions for themselves (Moe and Wilson 1994, 14).

Statute

Two groups vie for authority in determining the meaning of ambiguous statutory text: officials in the executive branch (“bureaucrats” or “agencies”) and judges. Stephenson (2006) argues that Congress, in legislating, consciously decides how to allocate power to each branch through the structure of its laws and its decision about whether to explicitly delegate policymaking authority to bureaucrats, but this vastly overestimates legislators’ ability and inclination to estimate the effects of the laws they pass in granular detail. It is difficult to say in practice exactly what kind of text would ensure a total allocation of decision-making authority to the executive branch or to the courts.

In modern administrative law, there seems to be a ready answer in Chevron: Step One says that clear language must govern, meaning courts should enforce it against reluctant bureaucrats; Step Two says that where textual meaning is ambiguous, an agency’s reasonable construction should receive deference from the courts. But several complications keep this seemingly straightforward rule from providing clear guidance. First, what does it mean for language to be clear or ambiguous? Whether some language is clear or not is often precisely the

48 See Section 1B. I don’t mean to imply whether or not Stephenson’s stylized approach is a useful simplification for his purposes, only that his model seems to poorly describe how legislators actually think in the majority of cases.
issue being disputed in cases that reach a court. Second, what constructions are “reasonable”? How is the Step Two inquiry meaningfully different from asking whether the judge believes the agency has complied with the statutory text?49 Finally, and most troublingly, when will courts actually invoke *Chevron* and adhere to its two-step analysis? This so-called *Chevron* Step Zero question (Sunstein 2006) calls into doubt whether even a perfectly clear doctrinal allocation of interpretive authority would necessarily provide a reliable real-world allocation to match.

Since *Chevron* leaves many open questions and there are no entirely reliable verbal formulae to ensure that one branch or the other will have full authority to interpret ambiguous language, there is a fierce struggle over most agencies’ structure, especially as they are created. Because an enacting Congress knows that it will leave many unresolved questions (often in the form of ambiguous language), various interests hope to ensure that future disputes are decided by their favored interpreter. Proponents of the so-called “congressional dominance hypothesis” (most famously including McCubbins and Schwartz 1984; McCubbins, Noll, and Weingast 1987, 1989) optimistically believed that Congress could graft onto its statutes administrative procedures that would ensure that favored interest group constituents would be able to steer future development of the statute in their favor. Some of these procedures create opportunities for legislative oversight (ex post political control), but others ensure a trajectory of future policymaking by helping to predetermine the impact of judicial resolution of ambiguous text (e.g. liberalized standing requirements and notice requirements). Critics of this school of thought doubt that administrative procedures actually produce the results that legislators hoped for (e.g. Balla 1998). Moe and Wilson (1994) argue that presidents retain prerogatives for unilateral

action analogous to residual decision rights, leading many presidents to focus on developing policy-steering capacity outside of the legislative process (see also Lewis 2003).

Apart from these positive debates about who does effectively determine the meaning of ambiguous statutory language, the normative question of who should be responsible remains. Here it is appropriate only to note how very many trees have given their lives for this question, and move on.\textsuperscript{50}

\textit{Constitution}

Who decides what the many ambiguous provisions of the constitution should mean in practice?\textsuperscript{51} As Section 2 briefly outlined, constitutional meaning will be determined in large part through a process of political contestation, and no part of this contestation is more fraught than the question of institutional authority. The executive branch and judiciary are again the principal claimants of power, but here Congress inevitably retains its relevance as an ongoing player merely by passing laws.\textsuperscript{52} As with statutes, there is a vast literature about who should have the power to determine constitutional meaning that I cannot even adequately canvass here.\textsuperscript{53}

\textbf{B. What principles should be used to resolve ambiguity?}

Saying a text is ambiguous is not the same as saying it is meaningless or incapable of offering guidance. The mantra of “just follow it” cannot guide actors responsible for giving

\textsuperscript{50} A few good starting places are Mashaw (1985), Vermeule (2006), and Elhauge (2008).
\textsuperscript{51} This question provides the organizing principle for one well-known textbook of constitutional interpretation, Murphy, Fleming, Barber, and Macedo (2008).
\textsuperscript{52} Theoretically speaking, there is no reason why Congress should not also be an equally powerful claimant of authority to determine the meaning of the constitution (see Pickerill 2004). Especially in recent times, however, Congress has increasingly adopted a passive role, in which legislators pass laws without much regard for constitutional sanction and then wait to see whether the Supreme Court accepts them as constitutional or not. Whittington (2007) provides a historical exposition of why this judicial supremacy has come to be accepted.\textsuperscript{53} Murphy, Fleming, Barber, and Macedo (2008) offers a good guide to these debates.
meaning to ambiguous text in the same way it purported to do for interpreters of clear text, but it is a mistake to imagine that the task of constructing ambiguous text is equivalent to “just making something up.” Llewellyn approached this question by reference to the “situational concepts” he saw at the heart of many rules, which describe an area of application but stop short of explicitly defining it. Using such concepts

is a technical, and verbal, device for giving almost self-operating satisfaction to the need for constant infusion into the technical structure of osmotic feeding from the social order; it is a technical, verbal device for getting that need filled imperceptibly, effectively, without men’s having to think about the need. (Llewellyn 2011, 107)

In other words, use of ambiguous, yet still descriptive, textual language allows interpreters to retain a legal mode of thought while necessarily having recourse to questions about how the (realized) rule fits in with the larger social order. Llewellyn recognized that such forms would necessarily discomfit people who believed that law should be reducible to clean logic—but himself urged that law should be thought of more expansively. When dealing with ambiguous texts, then, there can be faithful constructions, even recognizing that no one interpretation is uniquely faithful.

*Contract*

Contrary to what one might expect, contracting parties’ self-interest is not so strong as to prevent them from giving weight to common values as they seek to resolve disputes over ambiguous text. Rather than simply scrapping for advantage, parties often invoke what they see as relevant rules and precedents to support their side. Ideas about mutual accommodation are themselves norms, and often facilitate “reasoned elaborations” that are best understood as principle-based rather than simply as pragmatic compromises (Eisenberg 1976, 639-44, 649, 666).
There are several techniques that judges use to discern the practical implications of ambiguous statutory text, with seemingly never-ending arguments about which are (or is) most sound. Textualists or formalists insist that many seemingly ambiguous provisions are really not ambiguous if they are read within the context of the whole statute’s structure. Sophisticated textualists are at pains to distance themselves from any “crude literalism” which would pretend that nothing more than a trusty dictionary will be necessary to make sense of a law, but still insist that the text can resolve many questions. Even the staunchest textualist must admit that some statutory language really is ambiguous, at which point they must fall back on some other decision rule. (For example, Vermeule [2006] argues that textualist judges should defer to the constructions of executive branch officials when text is actually ambiguous.) Intentionalists argue that ambiguous text should be construed in the way that best comports with the intentions of the law’s framers. Interpreters are to engage in an imaginative reconstruction of legislators’ point of view (as informed by reliable evidence\textsuperscript{54}) and make of the ambiguous text whatever the framers would have wanted. Relatedly, purposivists may ask what purposes the law’s framers intended it to serve and argue for construing the text accordingly. Alternatively, purposivists might profess indifference to the law’s intended purpose, instead pragmatically asking what ends the law is now serving and how these can best be served in making sense of the ambiguous text.

These approaches hardly exhaust the field, and many scholars have wondered whether a less theoretically compelling account might nevertheless be a stronger and more descriptively

\textsuperscript{54} Just what evidence is reliable is a point of major contention. A debate about committee-prepared legislative histories has raged for decades, with strict textualists arguing that treating them as authoritative wrongly empowers committee staff, who will then abuse this power by preparing misleading reports, or more generally attacking the idea that intent ungrounded in properly passed statutory language should be authoritative. For arguments against reliance on legislative history, see Starr (1987), Easterbrook (1988), and Vermeule (2006, chap. 4); for defenses of the practice, see Eskridge (1990), Breyer (1992), Katzmann (2012); for a retrospective discussion, see Koby (1999).
accurate one. For example, working in a Llewellyn-inspired tradition, Farber (1992, 536-37) endorses a movement variously called “intuitionism,” “prudence,” or “practical reason” which would have judges account for: “1) the court’s sense of the situation; 2) the overall coherence of the legal system; 3) the presumed purpose of the statute; 4) the legislative history, at least if the statute is recent; 5) the statutory language (a factor of particular importance)” (punctuation changed and footnotes omitted). Through consideration of these factors, judges can situate situations within familiar legal categories and then go on to consider “the particular equities.”

Although most debates about what interpreters should do with ambiguous text focus squarely on judges, the question is even more relevant for executive branch officials charged with making sense of the text in the first instance (Strauss 1990). Although bureaucratic interpreters of the law could confine their inquiry into statutory meaning to the question of what the relevant judges would require should the issue end up in court, reading these tea leaves may be very difficult if existing precedents speak to the legal questions only tangentially. At the same time, since administrators have superior knowledge of how the law operates on a day-to-day basis, their attempt to determine the best construction of the law in light of the agency’s established practices may lead to a better practical realization of the law’s purposes than more strategically trying to anticipate what the courts would make of the law (Mashaw 2005).

Constitution

Debates about the proper way to approach seemingly ambiguous constitutional text closely parallel those in the statutory realm, although with a few twists. Originalists, like textualists, insist that there is less ambiguity than first meets the eye, and that the constitutional text alone is capable of resolving many questions if it is taken in its proper context. “Semantic
originalism,” as advanced by Solum (2008), argues that this meaning must be apparent in the text itself, while other originalists (Whittington 1999a) believe that the framers’ intent is definitive in fixing the meaning of constitutional language in spite of difficulties in pinning those intentions down historically. Purposivists tend to argue in somewhat different terms, frequently citing the constitution’s nature as a “living” document to justify interpretations tailored to modern needs that could not have occurred to the framers. As alluded to in the introduction, despite some four decades of debate between originalists and living constitutionalists, many scholars of constitutional law in recent years have begun to seek common ground, a point I turn to presently.

5. Formalists and their Opponents: Agreements and Disagreements

In contract law, formalists debate neoclassicists. In debates over statutory interpretation, textualists face off against intentionalists and purposivists. And every undergraduate constitutional law class teaches of the divide between originalists and living constitutionalists. Sunstein (1999, 638-39) helpfully lays out the basic commitments of formalism, which are more or less constant across each of these areas of the law. Formalists (including textualists and originalists) are for: 1) promoting compliance with formal rules in all cases, regardless of the apparent sensibility of doing so; 2) ensuring rule-bound law (even when applying the rule seems to make little sense); 3) constraining the discretion of judges. In pursuit of this third goal, formalists generally seek to deny courts four powers: 1) making exceptions; 2) allowing textual meaning to change over time; 3) invoking canons of construction to push statutes in favored directions; 4) invoking legislative purposes to push otherwise unambiguous words in favored directions. As Sunstein is quick to point out, very few scholars or judges are willing to wholly embrace all of these goals or their opposites, and so the debate about formalism tends to be a
matter of “how much” rather than “yes or no” (640). Indeed, as I show here, there is a core of propositions that most scholars now agree upon—with constitutional scholars leading the way in explicitly embracing their points of commonality.

A. Basic Agreements

The basic agreement encompasses the following propositions:

0. Some parts of textual instruments are clear, and some are not.
1. (a) Those actors bound by textual instruments must treat clear text as binding, although (b) there may be exceptions (as discussed in section 3C).
2. When confronting text that is ambiguous, institutional competencies should largely dictate which actors have the most control (or the last word) in determining the practical meaning of this text.

The first of these propositions I number as zero because calling it into doubt renders all the rest of the discussion extremely dubious. Few people believe that it will be clear which text is clear—indeed, just what constitutes clarity is one of the major sources of disagreement discussed below. Rather, to accept proposition 0 simply requires that one believe that clear text exists and is determinate, at least sometimes and in some contexts. Rejecting this implies that

55 Nevertheless, popular discourse about the law often does seem to embrace what Leiter (2010) dubs “vulgar formalism,” a particularly hardline version of formalism which insists that judges must always follow the law, period. Such a view seems to have a deep enough hold over American political discourse that our federal judicial nominees almost universally espouse this creed before the Senate Judiciary. Although there is plenty of reason to doubt their sincerity given the political and professional stakes, there is also reason to think that calling this point of view a “creed” is more than a casual metaphor. Smith (1999, 1066) argues that when lawyers engage in textual interpretation, their enterprise is closely analogous to the exegesis of scripture, where texts are “treated as if they speak in a perpetual present with a single voice.” He finds a paradox running through the modern enterprise of law: it runs on a certain kind of faith, but nearly all reflective and sophisticated observers (following the realist vision that Holmes elaborated) disavow that faith as anything more than metaphorical. (Tamanaha [2009] convincingly shows that this paradox pre-dated modern legal realism through examples of sophisticated scholars and practitioners of the 19th century who were able to see the shortcomings of thinking about law in strict formalist terms.) Smith believes that scholars hoping to fulfill the realist ambition of “understand[ing] law as it really operates” must not treat this faith as some embarrassing or “vulgar” superstition, but rather take it seriously as one of the crucial elements of the rule of law in our society.

This strikes me as a very important challenge to realist scholarship, especially to the extent that we hope to understand how lawyers and judges usually think. As I argue below, however, it seems to me that even government officials are able to publicly admit that their quasi-religious obedience to “the law” can provide more guidance in some situations than in others. They may, nevertheless, retain the language of formalism, and perhaps we should be cautious in condemning this as inaccurate if its use promotes the rule of law in some important ways.
law is at best a farce and at worst a sinister way to clothe the exercise of power in respectable
garb. This position is not a straw man. The more cynical and extreme versions of realism,
notably Jerome Frank’s *Law and the Modern Mind* (1930) and Roberto Unger’s *Critical Legal
Studies* (1986) seem quite close to embracing this position, although arguably neither does. Too,
certain judicial politics scholars often adopt rhetoric that suggests that law is little more than a
sideshow (e.g. Segal and Westerland 2011, which offers an affirmative answer to “Are Supreme
Court Justices Merely Legislators in Robes?”). For such arch-skeptics of the law, a belief in
law’s independent efficacy is superstition, albeit a widely held one. Perhaps there is little hope
for constructive dialogue between such people and those who see law in operation and believe it
matters. 56

Given a belief in the existence of clear and determinate legal text, nearly everyone
subscribes to 1a: legal texts are capable of clearly communicating some requirements, and
following the law demands giving force (both normative and practical) to these requirements.
Skeptical realists will insist that only a few rules are able to be fully determinate, not because of
“willful or corrupt administrators,” but rather because meaning is “read into, and through the
prisms of, the institutional structure of the going lawways and of the existing lawmen”
(Llewellyn 2011, 44). Nevertheless, at the limit, nearly everyone will agree that requirements
embodied in text can be clear enough to be effectively binding. 57 Although most rules, then, take

---

56 I doubt that anyone would argue against statement 0 from the other direction—that is, that all text is clear. Even
the staunches formalist must admit that some text is nonsense incapable of directing action.
57 Whether the historical legal realists would agree with this is not such an easy question. Although their detractors
frequently misunderstood their position on this point, realists did not doubt that law was somewhat determinate;
rather, they doubted that the outcomes of cases could be understood in terms of the formal legal rules cited by judges
rendering their decisions (see Leiter 2007, 25-30, 108). Most legal realist writing concerns areas of the common
law, rather than statutory law, so it is unclear what many realists would have thought about whether some statutory
text could be so clear as to be effectively binding. It seems that Llewellyn would have subscribed to such a belief, at
least for some limited set of statutory requirements: “It takes a hard rule, a well-made rule, compact, clean of
outline, definite and sure of scope and detail, and built to a feasible end, to pass through the institutional machine
and come out still its original self, still wholly master of its destiny and indeed of its meaning” (2011, 44). Schauer
on meaning partially from their context, clarity does make a difference, and the idea of textually
codified rules should not be thought of as fundamentally defective merely because they fall short
of full determinacy.58

In constitutional law, an emerging consensus around the possibility of clear, binding text
has led some to adopt the slogan, “We are all originalists now” (Solum 2011). Scholars of all
political stripes are increasingly likely to accept the idea that constitutional text sometimes bears
a fixed and definite meaning, and when it does this text should and can be taken as controlling.
To be sure, agreeing to this in principle does not mean that disputes about what the meaning is
disappear. As Whittington (1999a, 4) puts it, “[originalism] cannot be expected to free judges
from the exercise of contestable interpretive judgment.” This “new” originalism’s modesty
allows its wide acceptance, as those who espouse it are not committed to say that every
constitutional question is determinate—or that, for example, Justice Scalia’s views about the
original meaning of the second amendment are the correct ones.59

Although there has been not been a comparably public détente in either the contract law
or statutory interpretation literatures (perhaps because the rhetoric was never so heated in the
first place), it is worth emphasizing the existence of a common core of understanding in these
realms, as well. Clear text must control, or else there is little reason to bother with textual
instruments in the first place. At the same time, not all text is clear and not all questions are

---

58 “[Rules] deserve indeed more sympathetic and flexible study than it is easy to give them if one begins by
assuming that they exert, in bloc, the control which, in bloc, they officially purport to” (Llewellyn 2011, 49).
59 This modesty drives some defenders of “old” originalism and its bold promises to despair. For example,
criticizing Jack Balkin’s form of originalism, Alicea laments that “Balkin’s theory robs citizens of the ability to
condemn a decision on neutral grounds acceptable to all citizens,” i.e. by means of uncontestable historical reasons
(2010, p. 92). Such “neutral” versions of history, however, are likely to be elusive in practice, with competing
historical evidence in many cases creating ambiguities that no value-free theory can resolve.
answered in textual instruments. Ambiguities and silences ought to be addressed directly, putting aside fictions of complete contracts and statutes except in exceedingly simple cases.

Clear text is binding then—except when it isn’t. It seems that nearly everyone is willing to agree to 1b: exceptions exist. Contractual partners and government officials are rule-followers, but they are not only rule-followers, and sometimes their judgment should be able to displace the apparently clear requirements of a textual instrument. This may happen for any of the reasons noted in Section 3.C—including informal consensus among all affected parties that applying the law as written would not serve their interests. Such a consensus is most likely in the contractual sphere, where there are fewer parties. True consensus is less likely for statutory law, but sometimes government officials and affected interest groups may all believe that following the letter of the law fails to advance any important purpose and decide to adopt some work-around.

Conceding that the law will sometimes be put aside need not mean “we find ourselves cast upon the arbitrary action of officials—judges or others—as on a desolate and despairing reef” (Llewellyn 2011, 145). Rather, successfully policed social norms of judging may provide guidance and constraint; there are clear ideas about what it means to honorably execute the office of a judge or bureaucrat, and the men and women who come to occupy those offices seem to be deeply affected by those ideas.60 The responsibilities of office-holding include getting day-to-day operations of government to work smoothly, keeping the peace, and facilitating “stability in change” (Llewellyn 2011, 139).61 “Ground-level” officials will therefore be more devoted to “making things work” or “muddling through” than lawmakers delivering commands from on

60 For a treatment of the idea of an office that goes beyond law, see Heclo (2008).
61 Llewellyn (2011, 139) argues that “stability” and “change” are not rightly thought of as opposites in law; rather, the appropriate contrast to draw is between the opposing evils of “rigidity” and “erratic change.”
high, and as a result they will sometimes have to build structures that have only cantilevered support in a textual framework, so to speak. Forcing a statutory edifice to topple in the name of textual fidelity makes a mockery of the fact that the branches of government are nevertheless part of a single enterprise of governance (Siegel 2001, 325). Protecting the law’s reasonableness and coherence can be thought of as a key element of the rule of law—even if such action is rightly seen as being in tension with the values of predictability and transparency that are also key elements of rule of law (Dougherty 1994, 133-34).

Not everyone is equally sanguine about such attitudes, however, and disagreements about the appropriate scope of exceptions to clear text are discussed below.

Amongst students of interpretation of all stripes, there is also high-level agreement on statement 2: relative institutional competencies should dictate which actors have the most power to determine what to make of ambiguous text. As Section 2A explained, there is no single correct construction of ambiguous text, but this does not mean that anything goes. For text to remain meaningful (even if less-than-fully-determinate), constructions must be legally accountable, which means that some institutional actors must have the power to reject some constructions. Nearly everyone agrees that institutional competencies should determine who is able to make these determinations.

Unsurprisingly, formalists tend to think of competencies in terms of powers formally granted by statutes or constitutions. Because the legislature is responsible for making laws, its decisions about the law are binding and final—with a common-law judicial law-making function resolving only interstitial matters and bureaucrats deciding only those questions which have been expressly delegated to them. In practice, formalists often assume that judges are more likely to faithfully render the meaning of textual instruments than are bureaucrats (or parties to a
contract), which means that their judgments about constructions of ambiguous text should prevail.62

Those who are less committed to formalism are likely to combine some concern with formal competencies with a belief that practical competency should largely dictate which institutional actors should have the prerogative to construe ambiguous text. Although if a textual instrument is quite clear, formal competencies should govern, in the presence of ambiguity legal processes ought to consciously favor the most expert and objective decision-maker, as doing so should produce the best real-world results. Determining which actor actually fits this description in any particular case is a difficult problem—and quickly leads to significant disagreements.

B. Lines of Disagreement

For the core agreements outlined above, there are corresponding disagreements.

1. (a) What does it mean for text to have a clear meaning?
   (b) When should exceptions be desired or tolerated?

2. (a) What are the respective institutional competencies?
   (b) Should we come to a universal answer to this question, or case by case?

   In a 1989 law review article, Justice Scalia presciently ventured: “How clear is clear? It is here, if Chevron is not abandoned, that the future of agency interpretations of law will be fought” (Scalia 1989a, 520-1). This question bedevils far more than doctrinal analysis under Chevron, though. Without an answer to “how clear is clear,” it is impossible to say which parts of textual instruments fix meaning permanently and without any doubt (see Note 2005).

   In practice, formalists, textualists, and originalists have given one set of answers to this question, and their theoretical adversaries the opposite. With considerable oversimplification,

---

62 “Vulgar” formalists’ faith in judges’ fidelity to textual meaning seems, in the main, to be truly an article of faith. See footnote 54. Once this is called into doubt, one must make empirically grounded, realist arguments about why judges are likely to be more competent than other interpreters.
formalists have tended to argue that much text is clear enough to overcome reasonable doubts, so that law’s indeterminacy is a marginal rather than central phenomenon. Realists of various stripes have tended to be more skeptical of a drafters’ ability to formulate and verbalize their goals in ironclad language, and so see indeterminacy as pervasive.\(^6\)

Although the difference between formalists and their opponents about law’s ability to communicate are nearly always formulated in normative terms (“How much clarity should interpreters demand from text in order to see it as totally binding?), there are equally pressing empirical questions that are usually neglected. To learn about the ability of some words to communicate a particular concept in the context of normal language, we would almost certainly try to figure out how normal speakers would react to those words, without any question of how they “should” act creeping in. When thinking about how different actors will react to the words in various textual instruments, it is unsurprising that normative questions should demand attention—since we often argue about the appropriate roles for contractual partners or government actors when we discuss legal meanings. But this normativity should not cause us to neglect the empirical question of how much meaning words can convey as a matter of legal practice. “Ought implies can,” as realists like to point out, and arguably our best way of ascertaining what actors “can” do is to see what others in their position have done. Determined formalists will not balk at this, saying that a whole cohort of law-flouters does not make law-following any less practicable for the pure of heart; a personnel problem should not be confused with a conceptual impossibility. From a certain perspective, this attitude makes sense, but probably not from an institutional design perspective in which we imagine that future

\(^6\) Farber (1992, 534-35) spells out what I call a realist perspective in the debate: “Formalist interpretation, ultimately, relies on a faith in the raw power of the word to communicate, as if the perplexities of statutory interpretation were due merely to legal sophistry. Unfortunately, however, the need to understand context and purpose is inherent in language itself; it is not merely an invention of post-modernism.”
government officials will not be radically different than past ones. All this suggests that one of the most productive ways of advancing future debates about textual interpretation is to attempt systematic research about how binding specific textual instruments prove to be in practice. Contract lawyers presumably already attend to these questions in great detail so that they can ensure effective instruments. In the statutory realm, legislators should be given empirical evidence about the kinds of statutory language that tend to actually—in practice and not just as a matter of aspirational grammar—constrain bureaucrats. Chapter Three takes halting steps in this direction. More strategies for empirical research are badly needed, and could ultimately offer contract-writers and statutory drafters practical strategies for drafting clearer and more effective language.

The next area of disagreement has to do with question 1b: when should deviations from the text’s clear meaning be tolerated or desired? Formalists, textualists, and originalists believe that exceptions should be truly exceptional; others believe that textual requirements should be defeasible whenever compelling reasons for deviating exist. Taken to their logical extremes, both of these positions seem quite indefensible. Formalists unwilling to admit any exceptions privilege rule-following over all other values, notwithstanding any absurd, disastrous, or morally odious consequences that rule-following might entail. Few scholars, and even fewer practitioners, are willing to embrace this procrustean outlook. Anti-formalists placing clear textual requirements on equal footing with all other unspecified considerations destroy the commitment value that justified use of explicit text in the first place.

---

64 Justice Scalia, for example, famously calls himself a “faint-hearted originalist” and admits that there are many trap-doors to allow escape from awkward practical situations when following the original language clearly seems troubling. See Scalia (1989b).
Between these extremes, however, there are many reasonable positions. Sophisticated formalists, textualists, and originalists emphasize the importance of the ability to commit through text, for contracting parties, government, and for the popular sovereign. In his defense of originalism, for example, Whittington (1999a) emphasizes that fidelity to the intentions of the framers (as embodied in constitutional text) is the only way to allow citizens to effectively commit themselves, so that ignoring clear textual meaning—even ostensibly in the name of majority desires—ultimately weakens the popular sovereign’s ability to govern. Originalists and formalists who discourage exceptions also point out that textual instruments can be formally altered or abandoned if they fail to serve their desired purposes. Exercising these formal options preserves the fundamental ability to commit for the ultimate parties to the instrument (contracting partners, Congress, and the people), while interpreters who adopt ad hoc changes to meet exigencies arrogate that power for themselves (see Lawson 1997, p. 1831).

Two criticisms of this reasonable formalist position stand out. First, the formalists’ insistence that formal amendments to the text are possible does not always seem accurate. Especially in the American constitutional context, many question whether the super-majority requirements of Article V are so stringent as to make timely change impossible. Ackerman (1993) famously argued that the constitution was effectively, if informally, amended to support a stronger federal government during the New Deal, a sea change that achieved super-majority ratification as necessary support even if circumventing the super-majority requirements of Article V. One might similarly argue that putting aside clear constitutional meaning could be

---

65 Alternatively, originalists may offer a thinner argument offering only a theory of interpretation, rather than a full theory of how government actors ought to behave. Lawson (1997) uses an analogy to interpreting a recipe for fried chicken to emphasize this point, arguing that correctly understanding the recipe’s meaning is entirely distinct from asking whether the recipe is a good one worth following.

66 Levinson (1990) provides a good exploration of these issues.

67 Some might argue that this change happened exclusively through constructions of ambiguous and indeterminate statutory text rather than through creating exceptions to clear constitutional limitations. Others, of course, argue that
appropriate where there is little practical hope of reform through the Article V process, though
this would certainly call into question the Constitution’s constraining power. In the contractual
realm, if the transaction costs of including additional contractual terms are very low, it would
make sense for judges to have a stronger presumption that the written terms are exhaustive.
Harder-to-codify arrangements, on the other hand, would mean contractual text is better
understood as incomplete and possibly misleading, perhaps giving judges greater reason to put
aside (fairly) clear text (see Posner 1998, although this argument is a slight variation of his).

    In this dissertation, Chapters Four and Five provide examples in which courts proved
sympathetic to novel interpretations of laws in the absence of legislative action. The general
pattern that emerges seems to be that courts attempt to adhere to the meaning of (fairly) clear
legislative text that they believe its enacting coalition intended, but as laws become “stale” and
begin to seem unworkable or inadequate to some pressing need, judges expect the legislature to
“renegotiate” in good faith. Since Congress can “renegotiate” largely on its own terms68,
changing the law’s substance to account for changes in the body’s disposition, this seems to be a
reasonable requirement. When the legislature nevertheless remains on the sidelines, the court
may be more likely to permit a unilateral reworking on the part of the agency, allowing priorities
of muddling through to trump fidelity to statutory text.

    The second criticism of formalist opposition to exceptions proceeds from utilitarian
premises and asks whether there are any good empirical reasons to believe that consistent
application of formalist interpretive practices will produce better results. From this point of
view, formalism must be defended empirically, “pragmatically, with close reference to the likely

68 Of course, though, in the context of divided government the president’s veto creates a significant qualification to
this power.
performance of various institutions, and in terms of its consequences; it is not easily defended by reference to quasi-theological claims about the nature of law, legitimacy, or democracy” (Sunstein 1999, 641, 662-668).\(^{69}\) Formalists may simply reject the premise, but probably more would in principle be willing to engage it. There is an opportunity, then, for empirical investigations of the costs and benefits of exceptions to clear textual requirements to productively advance discussions between formalists and their adversaries.

Moving on to the disagreements surrounding 2a, about comparative institutional competencies in construing ambiguous text, the role of empirical evidence is potentially even greater. As discussed above, there is at least a limited generic agreement that institutional competencies ought to decide which actors are given the greatest authority in determining ambiguous text’s practical meaning. As with the question of when making exceptions is appropriate, some formalists may insist that these competencies are wholly a matter of formally designated authorities rather than practical capabilities, and there may be little possibility for constructive dialogue between those who adopt this premise and those who reject it.\(^{70}\) Plenty of other formalists, textualists, and originalists explicitly ground their arguments about who should

\(^{69}\) As footnote 54 points out, this dismissal of the formalist mind-set as “quasi-theological” may be rather too nonchalant; one can rationally wonder whether law as an enterprise might benefit from some degree of dogmatism not easily reducible to utilitarian considerations. In the statutory context to which Sunstein is referring, it may also be hard to “empirically” evaluate how much the commitment to letting the people decide issues through their representatives is worth, just as it is hard to evaluate the worth of the ability to prospectively contract or fix the limits of government through a process of constitutional deliberation. Even sharing Sunstein’s basic utilitarian orientation, then, one could hesitate to put too much weight on what empirical examination is capable of definitely resolving. For similar arguments, see Vermeule (1999).

\(^{70}\) Analogies across the different areas may be somewhat limited here. For example, Whittington (1999a) espouses a Thayerian belief in judicial deference where constitutional meaning is ambiguous, arguing that this response best honors constitutionalism’s principle of popular sovereignty by allowing elected representatives to decide questions that the constitutional document leaves open. But the counter-majoritarian difficulty motivating this perspective does not necessarily carry over to the statutory or contractual settings, where it is less clear why “whoever has occupied the field” of ambiguity first has any greater claim to legitimacy.
be responsible for determining the meaning of ambiguous text in terms of which institutional choice will yield the best practical results.

Generally speaking, if we think that judges are extremely competent and likely to get “right” answers when they are asked to resolve cases, then the dislike of judicial discretion that generally characterizes formalists will make little sense (Sunstein 1999, 642). In the contractual setting, if judges were perfect at evaluating evidence, we would want a “soft” parol evidence rule that would allow judges to weigh evidence of unwritten and relational elements of agreements and enforce those that were actually meant to bind the parties (Posner 1998, 543-44). On the other hand, if judges are less competent, there is less reason to trust their conclusions about what agreements were final and which were elements of the negotiating process. A belief that judges are “radically incompetent” leads to a sense that they should not engage in ambitious gap-filling, since their attempts to do so will likely misfire (Posner 2000). 71 Analogously, in the statutory setting judges’ constructions of ambiguous statutory text should be favored over bureaucrats’ only if there are reasons to believe that judges are better at making sense of the statute. Leading formalist accounts (especially Vermeule 2006) base their defense judicial deference where text is ambiguous precisely on a low estimate of what judges are capable of relative to actors in the executive branch. Anti-formalists (e.g. Eskridge 1994, Strauss 2005) emphasize judicial ability to adapt laws to changing contexts. In the constitutional sphere, it tends to be very difficult to say that any construction of ambiguous text is “best,” and so debates about institutional

71 Murray (2002), who has a fairly high estimation of judicial ability, offers a blistering attack on contract law formalists who assume judicial incompetence in weighing evidence: “Theorists seek to impose their idiosyncratic views on a process that is necessarily eclectic and administered by heterogeneous decision-makers who are largely committed to an unscientific and untheoretical methodology often called practical reasoning…. Judges using practical reason know that their engagement with the ‘surrounding circumstances’ does not rise to the level of systematic empirical inquiry because the parties have neither the time nor resources to provide it, and even with sufficient time and resources, complete data would be undiscoverable. The formalist notion that judges should not be permitted to consider necessarily incomplete empirical observations, however, suggests a monistic and dangerous approach” (908-909, footnote omitted).
competency often reduce to questions about the quality of deliberation each institution is likely to provide.\textsuperscript{72} Waldron (2001) and Kramer (2004) favor legislatures as the best representatives of “the people themselves”; Dworkin (1986) prefers courts; Posner and Vermeule (2010) are sanguine about leaving things to the president and executive branch.

Presenting empirical evidence about which institutional actors do the “best” job at resolving the meaning of ambiguous text—especially evidence convincing to overcome strong prior assumptions—clearly runs into several difficulties. Most obviously, there is no consensus about the meaning of “best.” Many political scientists go so far as to think of the overall policy-making process (from law-making to implementation to battles in court) as producing outcomes valuable only inasmuch as they match actors’ fixed “ideal points”—in other words, every policy decision is a political win for some and a political loss for others, and “optimal” policy is wholly relative. This outlook is clearly more relevant in some cases than others, however. As I argue in the dissertation’s Introduction, policy preferences are not always so solid or antagonistic, and a people from all kinds of institutional perspectives frequently talk about “making policies work,” or “doing things better.” Some will undoubtedly dismiss such talk as so much piffle, but I am more optimistic that it represents something real.\textsuperscript{73} As a result, I see empirical explorations of what kinds of legal language and legal interpretations lead to workable, sound policy as a promising research frontier.

Even if (relatively apolitical) empirical evidence about “good” policy can be had, though, there remains a deeper problem: who is going to evaluate this evidence? Political scientists and

\textsuperscript{72} Eskridge and Frickey (1994, 90) note that the Supreme Court is especially likely to “disrupt the equilibrium” achieved through the political process “when the national political branches have failed to deliberate on the relevant constitutional values.”

\textsuperscript{73} Judge Patricia Wald (1981, 145) confessed that judges are often badly in need of expert analysis of policy consequences absent from the legal record. For some issues, this might be resolved if judges could have an extra opportunity to call in the parties’ own experts—Judge Wald’s recommendation. In other cases, however, larger empirical issues would be better addressed through reliance on unaligned scholarship, if indeed such scholarship existed.
legal theorists tend not to trouble themselves with this question, concentrating on answering institutional design questions in the abstract. As a result, discussions of formalism, textualism, and originalism adopt a global perspective, with opponents in the debates trying to justify their favored approach as the correct one. This leads to question 2b: is there a point in trying to adjudicate the institutional competence question globally? While I do not flatter myself to think that many people engaged in the enterprise will care, I suggest that there is not. To illustrate the point, consider the empirically-minded defense of textualism offered by Vermeule (2006). Vermeule proceeds from the premise that judges are generally bad at figuring things out once they get beyond the plain meaning of statutory text, and consequently believes that judges should get out of the way of more knowledgeable executive branch actors except when they violate some very clear statutory requirement. This outlook may appeal to a few judges with an Olympian image of themselves as looking down at all the feckless mortal judges beneath them—Justice Scalia (not coincidentally, Vermeule’s old boss) amongst them. When addressing most judges, however, starting from a premise of global judicial incompetence will ensure a poor reception. Even if we could somehow present global evidence that judicial overruling of executive branch interpretations was, in general, harmful (and mounting such a case would surely be a hugely daunting enterprise, both logistically and conceptually—to say nothing of the claim’s veracity), it seems unlikely that those who wield power would be particularly interested in hearing about it.

---

74 Rakoff (2010) offers another argument to this effect, explaining why interpreters of all kinds are unlikely to find the right approach to understanding statutes through recourse to any “theory.”

75 Judge Posner (2008, 215-6) puts the point more acidly: “Either [Vermeule] has no insight into what persuades judges, or, more likely, though his book urges a radical change in judicial behavior, this is just a rhetorical trope and judges are not actually a part of his intended audience, which is limited to other professors, who delight in paradox. … Academics who are serious about wanting judges to change have to appeal to their self-interest. To tell judges, as Vermeule in effect does, that they are so dumb that they cannot even administer the absurdity exception to literal interpretation, and so should give it up, will not strike a responsive chord.”
There is the possibility for much more focused work, however. Empirical work on particular questions or particular kinds of statutory language could much more plausibly show contextual institutional advantages that might seem persuasive to policymakers confronting questions within that context. If one could show evidence to judges that their rulings consistently fouled up some area of the law, leaving it confusing, unworkable, and generally reviled, perhaps they would attempt to be more modest in that area. (Alternatively, institutional reformers might consider removing such questions from their jurisdiction.) By increasing the amount of context-sensitive empirical evidence available, we can also encourage an empirical style of thinking for judges as they determine how formalist an approach to take in a given case. Rather than asking, “Am I a formalist?” a judge could instead address the questions, “Do the institutional facts in this case justify a formalist approach, or are the theoretical conditions that would support a formalist approach absent here? In the latter case, how can I determine which actor is best situated to interpret the relevant statute?” This line of questions could continue:

Does the paper rule currently serve as a good approximation of the real rule? If not, what path has the evolution of the real rule taken? What are the concerns that have motivated changed interpretations of the textual instrument over time? Did unintended consequences arise from some aspect of the paper rule, forcing changes—or did the world change, making the original textual instrument at least partially obsolete? What unintended consequences did the new interpretation cause? Are the drafters of the original textual instrument (either specific people or institutional actors) likely to be able to formally revise the paper rule to adjust to and deliberately shape the ongoing evolution of the real rule?

In deciding how to orient themselves to a textual instrument’s requirements (however clear), judges—and bureaucrats and legislators, besides—must try to answer questions like these. In

---

76 Oversimplifying, some of the best work by empirical scholars of regulatory policy in the 1980s seemed to provide detailed evidence for such a proposition, e.g. Katzmann (1986), Melnick (1984), and Rabkin (1989).

77 Judge John Newman (1984) indicated that such questions frequently informed his decisions.
doing so, they face serious resource constraints. Often, the vast majority of the information available to them is presented by precisely those parties who are least objective about the matters concerned, because most self-interested. Chapters Four and Five provide extended discussions of how judges might already be engaging in this kind of thinking when interpreting very old statutes, thereby implicitly addressing what Calabresi (1982) called the problem of statutory obsolescence. The multi-institutional narratives I provide in those chapters—which are drawn in some substantial part from the historical reviews found in judicial opinions—are of the sort I believe would help judges (and bureaucrats and legislators) to form a sound empirical judgment about relative institutional capacities. There would still be plenty of room for disagreement about contextual competencies, and of course far deeper divisions rooted in competing political ideals are unlikely to disappear as government officials struggle over the proper construction of less-than-fully-determinate text.

To connect this discussion back to the Introduction, and end this chapter with a dramatic flourish, then: I believe that the inquiries envisioned in Section 5 are the future of a fruitful legal realism. Highlighting the difference between paper laws and real laws as these historically-minded accounts would do has the great virtue of aligning in many ways with how actors across the policymaking process already think about problems – in legislative histories, judicial statements of fact, and preambles to agency rulemakings. Both for academic analysts thinking about how law really works and for policymakers thinking about how their interpretations of legal text fits into particular contexts, expanding and systematizing this enterprise offers the best chance of moving beyond bitter doctrinal disputes and toward a more rational and effective law.

78 As I lay out in Chapter Four, disagreements along these lines seem to be one of the most salient empirical differences between the justices in the majority and the dissenters in Massachusetts v. EPA (2007).
Cases Cited

McCulloch v. Maryland, 17 U.S. 316 (1819)


Works Cited


Chapter Three – The Grammar of Legal Determinacy, Contestation, and Constraint

With the conceptual framework of Chapters One and Two now in place, the remainder of this dissertation goes on to flesh out the practical dynamics of interpreting and applying statutory laws by examining particular statutes. Throughout these chapters, I paint a portrait of law-as-policy-determinant, rather than law-as-legislative-process-output, and begin to show how we can usefully study statutes’ power to constrain government actions—including the limitations of that power, and the ways in which executive branch actors and judges relate to statutes that they believe provide insufficient or unacceptable guidance.

This focus often leads me to adopt the perspective of those who interpret and use the laws rather than those who write them. The principle benefit of this approach is that I am able to provide a treatment of law-in-action, in the realist spirit, that is badly underrepresented in contemporary political science. As I described in the Introduction, by doing so I hope to open a novel and fruitful research agenda that can potentially be of great use to lawmakers in understanding how they can write more effective laws.¹ The drawback of my focus is that treating the content of statutes as fixed threatens to make them seem like mysterious and immovable objects, resistant to rational understanding especially at the level of particular language. This sense of inscrutability may sometimes be realistic, in the sense that the law’s interpreters themselves may find the use of one phrase or wording instead of another hard to

¹ To judge a law as “effective” is distinct from judging the worth of the policies it instantiates. Determining what policies are good is, to a close approximation, the purpose of politics. Reframing the inquiry as a question of whether benefits exceed costs encourages certain styles of analysis, but it does not remove politics from the equation—determining the valuable and the costly once again requires fundamentally political judgments. In the largest sense, then, successful policies can only be defined as those that history judges kindly. But laws’ intermediary role can be judged independently, and it is that intermediary role of attempting to control government actors that I focus on here.
understand, so that language becomes a sort of brute fact that must be dealt with rather than a source of clear meaning to be understood.

From a different perspective, though, many of the forces shaping legislators’ choices about particular language are well-understood. Moe (1989) argues that choices about bureaucratic architecture are dictated largely by interest group politics, since interest groups are highly attentive to these choices and average voters are not. Where interest groups are unable to block legislative programs unfavorable to them but retain some bargaining leverage, they may be able to secure statutory features that limit the damage leading legislative compromises to produce statutory design choices wholly unrelated to technical policy demands. Sunstein (1995) explains that, in the process of securing compromise, obscurity may be an asset; legislators will often be willing to acquiesce to a murky provision leaving uncertainty where a clearer articulation of requirements would force them into active opposition. Fiorina (1989 [1977]) suggests that legislators will happily delegate important decisions to the bureaucracy to provide themselves a means of later intervening on behalf of their dissatisfied constituents. Less cynically, we could imagine that Congress might choose open-ended language to facilitate a learning process in the bureaucracy and courts that could precipitate legislative improvements in the future. The work in this dissertation cries out for investigations of how these legislative processes translate into particular choices about statutory language that could then be matched with studies of how the statutory language affected and constrained the formation of policy, and I hope to undertake such work in the future.

Without delving into legislative politics for now, in this chapter I provide a portrait of the constrained flexibility that statutes create for bureaucrats and how it plays out in practice. At least in the context of modern American regulatory statutes, laws are meant to both empower and
constrain government officials to effectuate the policies envisioned by legislators. Rather than completely dictating every policy choice, these laws are intended to create a workable balance for those executing the law. Too much flexibility leaves unelected bureaucrats with roving mandates to enact whatever policies they please. Too little flexibility prevents expert officials from utilizing the full extent of their knowledge and thus makes it difficult for them to produce viable regulations in a complex, dynamic world. Successfully striking this balance is the challenge at the heart of the modern administrative state, and in this chapter I explore the mechanisms by which it is negotiated amidst competing claims about what statutory fidelity requires. What does it look like in practice for an agency to run up against legal constraints, and how does this process work? What particular characteristics must statutory language have for Congress’s attempts to constrain agency discretion to be effective?

In asking these questions, I seek to problematize the idea of legal constraint that is so often taken as self-evident by political scientists. As the Introduction’s discussion of competing models explained, many political scientists treat statutes as if they create transparently clear requirements for those administering the law. Contemporary bureaucracy scholars generally model statutes as creating precisely bounded zones of discretion (e.g. Epstein and O’Halloran 1999; Huber and Shipan 2002), which they believe courts will be able to straightforwardly enforce against transgressing bureaucrats. Others have alternately lamented a complete lack of congressional control over a captured bureaucracy (e.g. Bernstein 1955; Kolko 1963; Lowi 1979; Schoenbrod 1993) or declared that Congress in fact gets everything it wants through clever use of ex ante controls embedded in statutes (e.g. Weingast and Moran 1983, Wood and Waterman 1993, McCubbins and Schwartz 1984, Lupia and McCubbins 1994, and McCubbins, Noll, and Weingast 1987, 1989; for criticism see Moe 1987). Some excellent recent scholarship has been
devoted to explaining how political context influences Congress’s use of several of these tools, including provisions for judicial review (Shipan 1997) and private attorney general powers (Farhang 2010).

Whereas these authors all make assumptions about how, and how effectively, law functions to constrain bureaucrats, this chapter seeks to illustrate constraint in action. In doing so, it shows that statutory language in action often confounds expectations. Like Balla (1998), who shows that inclusion of notice and comment procedures for setting Medicare physician payment rates advantaged constituencies nearly exactly opposed to those Congress had hoped to serve, my inquiry shows that congressional attempts to ensure constrained flexibility often give rise to unforeseen consequences in the hands of later interpreters.

My main argument is that constraint cannot be understood without close attention to the process of legal contestation. Legal constraints are only meaningful when some bureaucratic actions under a statute can be successfully challenged in court as impermissible, in spite of the justifications the bureaucrats offer. If executive branch interpreters can successfully fend off any legal challenge, then the statute ceases to be a meaningful legal constraint. As a result, understanding legal constraint in practice requires that we understand when legal contestation is likely to be successful. I explore which legal characteristics are likely to give rise to successful legal challenges based on the determinacy of legal language, and I make this discussion concrete through a quantitative and qualitative study of legal interpretations in rulemakings under the Clean Air Act (42 U.S.C. § 7401), one of America’s most far-reaching federal regulatory laws. As the Introduction explained, the descriptive quality of this empirical work allows us to focus on the what and how of legal constraint, making a vital contribution.²

² A different way to make this point is to say that this chapter advances our understanding by presenting insights into the mechanisms through which legal constraints are made real and then offers empirical evidence to illustrate the
Why Constrained Flexibility?

Why should Congress seek to write laws that deliver a balance of constraint and empowerment, rather than choosing one or the other value? Certainly not because doing so is easier. Writing laws that provide only constraint, or only flexibility, is relatively easy as a matter of drafting. Both can be phrased quite tersely. Laws embodying only constraint are generally not recognized as “delegating” any authority to executive branch actors at all: “Action X is hereby prohibited, and shall be punished by penalty Y,” is a familiar legal formulation that leaves little role for executive interpretation (though the importance of discretion in enforcement is often underestimated; Davis 1969). On the other hand, writing laws that leave things almost entirely to the discretion of the policymaker is also fairly easy: “The Administrator shall prescribe rules concerning Z that he determines will serve the public good.”

Laws of the latter sort—all flexibility and no constraint—are rarities in a society that values the rule of law. Though ostensibly working through “law,” in fact this statute would simply invite the arbitrary exercise of power, unchecked by any principle but political expediency. To the extent that an agency given such a vague mandate is constrained, the constraint will be entirely political rather than legal.3 There are, nevertheless, times in which Congress is willing to enact laws of this sort. Faced with what seemed to be an imminent threat of economic meltdown in October 2008, Congress hurriedly enacted the Emergency Economic Stabilization Act (Pub. L. No. 110-343, better known as TARP), § 101(a)(3)(C) of which stated: “The [Treasury] Secretary is authorized to take such actions as the Secretary deems necessary to carry out the authorities in this Act, including, without limitation…” It then listed some

---

3 In the terms of Whittington (1999) discussed in Chapter Two, executing a law such as this one is entirely a matter of construction rather than interpretation.
suggestions, but unusually left the door open for the Secretary to pursue almost any other action he believed was “necessary.” This exceptional statute shows that in dire circumstances legislators may care more about empowering the government to act than about subjecting its actions to legal constraints, notwithstanding their deep commitment to legislative supremacy.\footnote{Wallach (2010) more deeply considers the relationship of such legislative abdications to the rule of law.}

On the other hand, laws which explicitly admit little flexibility seem to conform easily to rule of law values, but in fact pose serious practical difficulties that make them rare. Where a completely clean linguistic formulation can transmit the purposes of an enacting coalition, that coalition would probably choose it. If the purposes a law is meant to serve are at all complex, however—and especially if the law is meant to balance values in tension with each other—such neat determinacy may be quite undesirable. Just as private parties may enter into incomplete contracts because of bounded rationality and the transaction costs of seeking completeness (Williamson 1989; also see Chapter Two), legislators forging a compromise are likely to favor some flexibility for those tasked with executing the law. Achieving democratic control and perfect predictability through precise and exacting laws may well jeopardize competing rule of law values, such as the reasonableness and coherence of the law, if it forces rigidity in response to changing or unexpected conditions. As a result, even those laws which we might think are meant to enact uncomplicated legal restrictions tend not to embrace constraint entirely to the exclusion of flexibility. The federal minimum wage (29 U.S.C. § 206 et seq.), for example, leaves the Secretary of Labor discretion to grant exemptions of many kinds, as do the laws criminalizing possession of controlled substances (21 U.S.C. § 816 et seq.).

In practice, then, we should expect to see few laws that confer total flexibility or that leave no room for maneuvering in implementation. When Congress chooses legal controls in its statutes, it seeks to establish a particular mix of constraint and flexibility. Given this
understanding, I now turn to considering the kinds of statutory language that push things toward constraint and show how this language is converted into effective constraint.

**Legal Determinacy and Contested Interpretations**

How do laws constrain—that is, how are laws made effectively binding on government actors? As discussed in the Introduction, many scholars skip this “how” question and skip directly to one of two polar opposite conclusions: that law constrains mechanically, “because it is law,” or that law constrains not at all, with ostensible disagreements about law merely masking an underlying political struggle. As I sought to establish in Chapter Two, the truth lies comfortably between these poles: law is both indeterminate and binding. The precise nature of its constraints is not simply a function of the text enacted by legislatures. Rather, the text as it is set out shapes the possibilities for government actors, whose interpretations of the law test its practical limits. This testing occurs through a process of contestation.

Although it seems patently obvious, the clearest implication of viewing legal constraint as a function of legal contestation is that if there is nobody to contest the executive branch’s interpretations, then there will be no effective legal constraint. This is not the same as saying that the law will be without importance; law could shape the way government officials conceptualize their professional roles, or it could provide a focal point for political expectations and thus serve as a touchstone for political constraints, such as those enforced through the congressional appropriations and oversight processes or through presidential supervision. But unless there is a way to legally challenge the government’s choices under a law, it is hard to say
that the law legally constrains government action. For example, § 108 of the Humphrey-Hawkins Full Employment Act of 1978 (Pub. L. No. 95-523) amended the monetary policy mandate of the Board of Governors of the Federal Reserve to include maintenance of economic growth in addition to price stability. Today, many commentators lament that the Fed seems to be ignoring its growth mandate in favor of its anti-inflation one to the detriment of the real economy, but whatever the merits of this position may be, the critics have no meaningful possibility of pursuing a legal case against the Board of Governors that would find the Board in violation of its legal responsibilities and enforce some remedy against it. The law affects political dynamics—the Fed would never openly flout either part of its mandate, and must explain its choices to Congress with reference to its dual mandate—but does not put in place a true legal constraint.

Many laws do afford the opportunity to bring legal challenges, however. When does this contestation occur, and what does it settle? At the highest level of generality, private parties will contest legal interpretations when they expect that doing so will likely benefit them: that is, when the probability-weighted benefits of winning a challenge exceed the costs of challenging \( p_B - C > 0 \). Challenges are therefore more likely when the chances of winning are high, the potential benefits are large, and the costs of pursuing a challenge are low. One clear implication is that challenges are more likely in policy areas where the stakes are large. Where regulations

---

5 Note that this is quite far from the colloquial meaning intended by calling a law “unenforceable.” Such statements nearly always refer to the practical difficulties faced by the government in making cases against violators of a law, rather than questions about legal limits on the government.

6 Recently, two notable administrative law scholars have called into doubt the very idea of legal constraints that provide effective checks on executive branch actions independent of any political constraints. Criticizing what they see as an anachronistic “liberal legalism,” Posner and Vermeule (2010) sweepingly argue that the modern executive branch possesses such advantages of expertise and decisiveness that it can easily maneuver around whatever legal barriers are created. The force of this claim rests on empirical questions: to what extent can and do people bring claims against government actions, and are courts willing and able to provide effective remedies if they find that executive branch actors have overstepped their authority? The empirical evidence provided in this chapter, while certainly not broad enough to provide global answers to these questions, suggests that Posner and Vermeule’s claims about the inefficacy of legal constraints are probably somewhat hyperbolic.
have the potential to require businesses to spend many millions or billions of dollars, for instance, challenges that would mitigate that burden if successful are likely. Another implication, developed by Farhang (2010) in a somewhat different context, is that the amount of litigation can be influenced by provisions altering the costs and benefits for potential litigants, for example by creating the possibility for treble damages, or recovery of litigation costs.

My main concern in this chapter is to discuss how statutory language’s level of determinacy affects the chances of affected interests bringing a successful challenge. I will turn to what makes language determinate in a moment, but the rough idea is that more determinate language is more precise in delineating what government action must consist of; making language more determinate pushes toward constraint, and away from flexibility.

What sort of laws will be most likely to support successful legal challenges of executive branch interpretations? The only really sensible answer is: it depends on the nature of the interpretation! After all, if there is a precisely determinate law: “The administrator shall perform actions 1, 2, and 3, and is forbidden from all other actions,” a legal challenge is likely to fare very differently if the agency cites this law as its basis for action 1 from when it cites the law as its basis for action 4. In the former case, the prospects of mounting a successful challenge should be approximately zero; in the latter case, near certainty. Predicting the effect of determinacy thus requires we also keep in mind the interpretive posture adopted by the agency, which could range from cautious to very aggressive (often described as “pushing the envelope”).

If we think about law in stylized terms as either completely determinate or extremely indeterminate, and combine them with interpretive postures, it ought to be easy to figure out when challenges would be successful. If the law is determinate, and the agency’s interpretation fell clearly within its meaning, challenges would be sure to fail—and would not be brought. If
the law is determinate and the agency’s interpretation represented a clear deviation, a challenge would certainly succeed. If this were indeed transparent, and there was someone positioned to bring a challenge, our expectation is that we would see no such deviations by the agency in the first place, since making such a vulnerable interpretation would simply be a waste of resources.

Now suppose that the law is highly indeterminate. (Supposing complete indeterminacy would render this exercise fruitless, equivalent to imagining an absence of law.) Given high indeterminacy, it will be harder to characterize the agency’s interpretation as either cautious or aggressive, but perhaps not impossible. Given a modest interpretation, mounting a successful legal challenge will be extremely difficult—the statute’s requirements are unclear, but on the face of it the agency’s interpretation seems to be a reasonable rendering. As a result, challenges should be rare. Given an aggressive interpretation, though, there may be somewhat more room for a challenge. Though a litigant contesting an agency interpretation will not be able to point to any statutory provision that has been unambiguously violated, they may be able to mount a case as to why the agency’s construal of an admittedly unclear text is unacceptable, and sympathetic judges might accept this argument. Table 1 schematically summarizes the preceding discussion.

Table 1 – How Legal Determinacy and Interpretive Posture interact to produce legal conflict, assuming dichotomous conditions

<table>
<thead>
<tr>
<th>Level of Determinacy</th>
<th>Cautious</th>
<th>Aggressive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete</td>
<td>All challenges fail → no challenges</td>
<td>All challenges succeed → agency avoids these interpretations → no challenges</td>
</tr>
<tr>
<td>Very low (highly indeterminate)</td>
<td>Challenges likely to fail → rare challenges</td>
<td>Challenges may succeed → frequent challenges</td>
</tr>
</tbody>
</table>
Darker shading indicates a higher probability of observed legal conflict

If we were to end our analysis here, we would think that challenges are made more likely by increasing indeterminacy, and by more aggressive interpretations. But moving toward more realistic conditions gives us reason to question these insights.

As previously discussed, real laws are unlikely to exhibit either complete determinacy or extreme indeterminacy. Instead, suppose determinacy is on a continuum, which I break into three categories (High, Medium, and Low) for ease of exposition. Given this added complexity, let us also assume that interpretive posture could vary more, and we will break this spectrum into three categories as well (Cautious, Medium, Aggressive).

If the law is highly determinate, things will not be so different than in the hypothetical case of the completely determinate law. A cautious interpretation will be nearly immune to challenges, which therefore won’t be pursued, and aggressive interpretations will be nearly impossible to defend against challenges, suggesting that the agency is unlikely to adopt such an interpretation in the first place. If the agency does make such an interpretation, it will need to do its best to diffuse legal challenges out of court. With high determinacy, the middle interpretive stance category is hard to make sense of.

If the law has low determinacy, things will not be so different than the case of extreme indeterminacy, either. If judges are somewhat deferential to agency judgments—generally a safe bet in the age of *Chevron* deference[^7]—challengers will have a hard time displacing cautious interpretations, making legal challenges rare. The most aggressive interpretations are likely to engender more conflict: the law in question is fairly vague, but even so the agency’s actions seem to be testing the limits of its legal powers and thus opening the door to legal contestation.

[^7]: The seminal case of *Chevron v. NRDC* (1984) stands for the proposition that courts are to defer to a reasonable agency interpretation unless it directly conflicts with statutory language. See Chapter Two for a more extensive discussion of *Chevron*.
A moderate interpretive stance should lead to a middling amount of controversy—it will be harder to establish that the agency’s interpretation is impermissible, but probably not impossible, making challenges worth pursuing in some cases.

Intuitively, laws of medium determinacy should lead to levels of legal contestation between those of high and low determinacy, but I argue that this is not the case. As with the other two categories, if the agency adopts a cautious interpretation, pursuing a challenge successfully should be quite difficult, and challenges should be rare. With a medium determinate law and a medium interpretive posture, however, challenges should be more prevalent. As its interpretations become less obvious, they are increasingly open to challenges; at the same time, those challenges are not so likely to succeed as to deter the agency’s choice in the first place.

Finally, a medium determinate law and an aggressive interpreter should lead to the most conflicts of all: the language seems clearly meant to be constraining the agency, but the agency’s interpretation seems almost to vitiate that constraint, making challenges fairly promising in spite of the lack of a clear-cut violation of the statute. Because it is difficult to predict how the reviewing court will receive the agency’s interpretations, pursuing outcomes all the way through the legal process becomes most necessary. This last insight is similar to Priest and Klein’s (1984) observation that rates of success for plaintiffs or appellants should tend toward fifty percent, regardless of the substantive legal standard involved. The logic is that litigants will look to settle only when it is fairly difficult to estimate which side is likely to win. If both sides can see a likely winner, they can come to a compromise approximating that end outcome with only minor concessions to the likely loser and thereby save themselves further costs of litigating. This prediction should hold, “other things being equal”—which they may well not be if the outcome
of the trial is far more important to one side than the other or the parties have an unequal
tolerance for risk.

Once again, we can summarize the discussion schematically (see Table 2).

**Table 2 – How Legal Determinacy and Interpretive Posture interact to produce legal conflict, assuming trichotomous conditions**

<table>
<thead>
<tr>
<th>Level of Determinacy</th>
<th>Interpreive Posture</th>
<th>Cautious</th>
<th>Medium</th>
<th>Aggressive</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td></td>
<td>All challenges fail → no challenges</td>
<td>(conceptually unclear)</td>
<td>Challenges likely to succeed → agency mostly avoids these interpretations → rare challenges</td>
</tr>
<tr>
<td>Medium</td>
<td></td>
<td>Nearly all challenges fail → rare challenges</td>
<td>Challenges have fair chance → frequent challenges</td>
<td>Challenges have good chance → very frequent challenges</td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td>Nearly all challenges fail → rare challenges</td>
<td>Challenges have slim chance → some challenges</td>
<td>Challenges have fair chance → frequent challenges</td>
</tr>
</tbody>
</table>

Darker shading indicates a higher probability of observed legal conflict

Rather than a monotonic relationship between determinacy and observed conflict, we see a peak of observed conflict in the medium determinacy category.

What are the implications of this for law as constraint? One can argue that even if the relationship between determinacy and observed conflict has been shown to be complex, there is still a simple monotonic relationship between greater determinacy and greater constraint. After all, in the case of high determinacy, conflict is rare because the agency is effectively constrained to hew to the law’s core meaning, and in the case of low determinacy, conflict is less common because it is difficult for challengers to show that the agency’s interpretation is impermissible, meaning constraint is low. The medium determinacy cases fall between these: constraint
happens sometimes. This is a fair point, but it hardly negates the importance of seeing when legal conflict arises and when it will be decisive in shaping legal meaning.

The theorized relationship between determinacy and constraint just laid out has important implications for the institutional distribution of power. Congress minimizes the need for later interpreters to supply its laws with meaning if it writes highly determinate statutes. It leaves the door open for bureaucrats’ interpretations to fix statutory meaning if it writes highly indeterminate laws (especially in an age of *Chevron* deference). Finally, it invests judges with the most authority when it uses medium-determinate statutory formulae, since the precise meaning of the law’s constraints is most likely to be elaborated through litigated conflict. One could argue that courts are a sort of “half-Congress, half-bureaucracy” condition, since judges are likely to try to balance legislative purpose against respect for agency judgments. In this case, a basically linear relationship would emerge: more determinacy means more control for Congress, less determinacy means more for the bureaucracy. Given the complexity of statutes and judges’ ability to inject concerns never anticipated by enacting legislators, treating the outcomes of legal contestation with as a sort of average between the enacting Congress and current bureaucracy’s concerns is a dubious assumption, though (see Katzmann 1986; Melnick 1983, 1994). Among other effects, channeling the search for statutory meaning through the judiciary will transmute political contestation into legal contestation—a transformation that is seldom neutral in its effects (Silverstein 2009). I return to these questions later in this chapter.

Given this institutional dynamic, we could think of legislators as consciously facing the choice of which institutional actors to empower when they draft new laws. For example, Stephenson (2006) models legislators as facing a conscious choice of delegating to agencies or courts. As noted above, however, legislators’ real choices about determinacy might be motivated
by political factors, especially the need to secure compromise through the use of vague language, rather than by beliefs about the optimal statutory language. Additionally, as examples under the Clean Air Act (CAA) will show, language intended to be highly determinate may nevertheless prove open to creative interpretations.

Analyzing the extent to which legislators will be able to purposefully and effectively choose levels of determinacy requires us to finally turn to the concept of what makes a law more determinate. Determinacy is intuitive enough: high determinacy is a function of language strongly and explicitly requiring non-discretionary, precisely specified, well defined actions without inclusion of any language suggesting that government officials have leeway in carrying the law into effect. Again, laws of the highest determinacy are unlikely to be viewed as delegating any policymaking authority at all—they would comprehensively define the policy’s “who, what, where, why, when, and how,” even if they left many details to be worked out during implementation.\(^8\) Somewhat less determinate are laws which require taking some action, but specify and define the criteria to be used in choosing the content of that action rather than precisely delineating what is required. Less determinate still are laws that confer discretionary powers (“the Administrator may…”) while specifying criteria for deciding when to use them (“if he finds that…”). Insertion of phrases such as “as the Administrator deems necessary” or “in the Administrator’s judgment” make laws less determinate.\(^9\)

---

\(^8\) In §202(g) of the CAA, for example, Congress legislated emission standards for light-duty trucks and other light-duty vehicles, precisely setting out which chemicals would be covered, what the numeric emissions standards would be, and what the implementation schedule would be. To be sure, this leaves questions to be resolved regarding measurement techniques, procedures for determining compliance, and responses to noncompliance—and commenters had plenty to say about the EPA’s choices on these subjects (EPA 1991). But this subsection of the act left little room for serious legal conflict, regardless of how any vehicle manufacturers may have felt about the policy it instantiated.

\(^9\) An example is §112(h)(1) of the CAA, which allows the Administrator to use an alternate work-practice-based standard if, “in the Administrator’s judgment,” a technology-based standard would not be feasible. Judgment about feasibility can certainly be contested with reference to objective facts, but the statute unquestionably pushes toward flexibility and discretion for the agency through use of this language.
A literary passage with a rich history of citations in legal writing\textsuperscript{10} points the way to a more general insight into determinacy. It is the dialogue on the meaning of words between Lewis Carroll’s Alice and Humpty Dumpty:

‘When I use a word,’ Humpty Dumpty said in rather a scornful tone, ‘it means just what I choose it to mean—neither more nor less.’

‘The question is,’ said Alice, ‘whether you can make words mean so many different things.’

‘The question is,’ said Humpty Dumpty, ‘which is to be master—that’s all.’

The quotation usually ends there—used simply to chide some interpreter for availing himself of “Humpty Dumpty’s facile prerogative over words” (Corwin 1934, 315). But I find Humpty Dumpty’s next remark more illuminating:

‘They’ve a temper, some of them—particularly verbs, they’re the proudest—adjectives you can do anything with, but not verbs—however, I can manage the whole lot of them!”\textsuperscript{11}

Humpty Dumpty is quite right: if a statute requires that an administrator “shall promulgate” a rule, there can be little question that promulgation must occur if the administrator is to be compliant with the law. Interpreting adjectives generally requires (and naturally allows) the use of more interpretive imagination than does following the commands of these “proud” verbs, and yet a handful of open-ended adjectives (and their adverbial forms) are staples of legislative drafting: adequate, appropriate, expeditious, feasible, necessary, practicable, reasonable, significant, sufficient. In few contexts is the meaning of such words self-evident. Undefined nouns, especially vague ones such as “public welfare,” can be just as troublesome.

It would be wrong to jump to the conclusion that such words are mere empty vessels, to be made to mean whatever the interpreter wants, Humpty-Dumpty-like. If this were so, no court would ever find reason to reject any interpretation of these words, but in fact they do. In the spirit of legal realism, and an understanding of the law as “the prophecies of what courts will do

\textsuperscript{10} Potter (2006) traces the history of judicial citations to Lewis Carroll more generally. Google Scholar finds 278 hits for this passage in legal journal articles and judicial opinions.

\textsuperscript{11} The dialogue is drawn from Carroll (1960 [1871], 268-269).
in fact, and nothing more pretentious” (Holmes 1897), legal regularities in the treatment of these words come to give them real meaning in their particular contexts. If courts come to regard the phrase “adequate to protect public health” as containing certain requirements, then the phrase is imbued with that meaning, and should provide a real constraint on the permissible choices for future interpreters (see Chapter Two, Section 2).

Nevertheless, depending on such words or phrases does not produce highly determinate statutes. As these phrases are taken up for the first time in new contexts, it may be hard to predict the extent to which judges will be willing to interpret them as creating binding constraints on bureaucratic agencies. Though a sort of common law of these provisions may be worked out through incremental decisions over time, it will be hard to predict exactly what these precedents will be at the outset. Consequently, laws whose pivotal phrases depend on adjectives, adverbs, or vague nouns are quintessential exemplars of medium determinacy. By the logic articulated above, we should expect laws depending on adjectival formulations to engender a greater amount of legal conflict than provisions built solely on sterner verbs.

The discussion to this point has been mostly abstract. To show the relationship of determinacy, contestation, and constraint in practice, I now turn to evidence from one particular corner of statutory interpretation: rulemaking under the Clean Air Act.

**Evidence from Clean Air Act Rulemakings**

If we want to understand the concrete processes by which statutes’ “paper” constraints are made real, we must observe these processes in action in some systematic way. I do so here by analyzing an original dataset of rulemakings under the Clean Air Act as well as delving into a few important sections of the law more deeply. This represents a promising (which is not to say
uniquely optimal) object of study for several reasons. First, agency rulemaking is one of the central mechanisms by which laws are given practical meaning in the modern American administrative state. This site of legal interpretation and policymaking offers some of the most important opportunities for interest group and democratic participation through the now-nearly-ubiquitous practice of notice-and-comment rulemaking—and thus a forum for some of the most important conflicts in legal interpretation of regulatory statutes. A small but promising literature is beginning to investigate trends in notice-and-comment rulemaking\textsuperscript{12}, but none of these works have attempted to understand the ways in which laws constrain rule makers. The empirical work presented here thus represents an important original contribution.

Within the universe of rulemaking, choices under the Clean Air Act (CAA) offer a fertile testing ground for exploring the nature of legal constraints for several reasons. The law is one of the most far-reaching and important statutes in the United States Code, especially after the major amendments it received in 1990. Environmental Protection Agency (EPA) rules under the statute govern the chemical content of the gasoline in our cars and the propellants in our spray cans; the emissions performance requirements for planes, trains, and automobiles, as well as motorcycles, motorboats, snowmobiles, lawnmowers, and industrial engines; and the allowable emissions from power plants, waste incinerators, and factories of all kinds. More rulemakings are proposed under the CAA in a given year than under any other statute, representing a very large part of the EPA’s workload.\textsuperscript{13} The CAA also offers a good balance of variation and

\textsuperscript{12} O’Connell (2008) and Gersen and O’Connell (2008) investigate macro-level trends in the amount and duration of rulemaking efforts in different political contexts. Webb Yackee and Webb Yackee (2006), Webb Yackee (2006), McKay and Webb Yackee (2007), and Naughton et al. (2009) document several aspects of the notice and comment process, providing evidence in support of such intuitive propositions as “the squeaky wheel gets the grease” (showing up and making one’s voice heard is half the battle), business commenters tend to be especially influential, and early participants can exert agenda-shaping power.

\textsuperscript{13} The choice to look at just one type of agency activity (rulemaking) while excluding others, such as agency guidance documents, internal allocations of resources, or choices not to act, unquestionably limits the generalizability of the findings here. First, it is possible that agencies will choose their action’s form based on its

133
stability in statutory language. The law is broad enough as to essentially contain many smaller laws, which share enough in common to support comparisons while still providing considerable variation in terms of the kinds of statutory language they utilize. Sections of the act provide stable points of comparison since the law has been only marginally amended since its bipartisan overhaul in 1990—and so this chapter’s large-n analysis is restricted to the years 1991-2009.

Several procedural aspects of the CAA also make it relatively attractive to study legal challenges to agency interpretations. Section 307 of the Act liberalizes judicial review of rulemaking, allowing any interested citizen to bring a petition for review of the rulemaking directly to the Circuit Court of Appeals (the idea being that such challenges will be exclusively legal in character, thus not requiring any fact-finding court). In the case of state-specific CAA actions, the appropriate local circuit has jurisdiction, while in nationwide CAA actions the D.C. Circuit has exclusive jurisdiction. By restricting its focus to CAA rulemakings with nationwide applicability, this study’s analysis thus has the benefit of dealing with consistent procedures and a consistent legal venue. In addition, section 113(g) of the CAA requires the EPA to publish a

substance (Tiller and Spiller 1999). Raso (2010) counters with preliminary evidence suggesting that strategic considerations do not drive choice of policymaking form, but courts are themselves attentive to this possibility, as evidenced by the D.C. Circuit’s ruling in Appalachian Power Co. v. EPA (2000, 1020, 1023), in which it struck down an EPA action undertaken through guidance documents and reprimanded the agency for falsely characterizing a substantive policy choice as not being a “final action” subject to the procedural requirements of the Administrative Procedures Act and the CAA. Overall, the stringency of the procedural requirements embedded in the CAA (and the courts’ willingness to make these constraints binding) somewhat mitigate this concern.

Second, apart from agencies’ strategic “form-shopping”, there is the question of whether rulemaking tends to arise from statutes of low determinacy. After all, intuitively, a very highly determinate law should just set out the course of action for an agency to follow rather than necessitating agency rule-writing. There is something to this claim, but we should remember that in practice almost no law is so determinate as to completely crowd out agency decision-making. As we shall see, the many sections of the CAA which engender rulemaking exhibit a great deal of variation in the levels of determinacy.

In short, the portrait of the legal constraints achieved by observing the rulemaking process under the CAA cannot be treated as a representative stand-in for all processes of legal constraint—or even all important legal constraints under the CAA. (The case of Massachusetts v. EPA [2007], discussed at length in Chapter Four, provides an example of one notable omission; the EPA’s discretion as to whether to regulate greenhouse gas emissions was challenged, and the Supreme Court held that the law’s constraints required the EPA to regulate. This case falls outside of this chapter’s data because it was rooted in a decision not to regulate, rather than in a rulemaking proceeding.) But this limitation does little to diminish the usefulness of the observational vantage point for present purposes.
notice of all settlements, which facilitates reliable and complete tracking of legal disputes even if they do not ultimately go to trial.

While a full discussion of the construction of the dataset is reserved for the chapter’s Methodological Appendix, I offer a quick description here. The unit of analysis is the “rulemaking chain”—a proposed rule, final rule, and any associated legal challenges to that final rule. Litigation is tracked to its conclusion, whether settlement out of court or an opinion by the D.C. Circuit. For each of these steps, data is recorded and (where possible) coded about the substance of the rule, its legal origin and justification, whether it met with adverse comments, and how the agency responded to those comments. For the purposes of this discussion, the most important piece of information is whether, and to what extent, the agency was convinced to abandon the legal interpretation embodied in its proposed rule in favor of some alternative interpretation advanced by commenters, possibly coupled with the argument that the agency’s proposed interpretation was impermissible. The EPA proposed 810 rules with nationwide applicability under the CAA from 1991-2009, some 42.6 per year. Of these, 146 (18.0%) led to litigation. The dataset I compiled includes 331 of these rulemakings—all of those that involved litigated challenges, and a sample of 185 (27.8%) of the rulemakings that did not lead to litigation. This strategy of stratified sampling is explained and defended in the Methodological Appendix. Because sampling was used, descriptive references to non-litigation-related aspects of the rulemakings, including the presence and influence of legal objections, reflect sample-weighted estimates of means rather than observed quantities unless otherwise noted.

Two factors slightly complicate the search for legal constraint within the notice-and-comment rulemaking process. First, not every instance in which the agency responds to comments on its proposal by changing its policies is attributable to legal constraint—often the
agency may simply be persuaded of the merit of an adverse commenter’s position. To attempt to distinguish instances where the agency worried about having its interpretations overturned in court from those in which commenters persuaded regulators, I code whether commenters raise legal difficulties with the agency’s interpretation, mostly by seeing whether they make reference to the underlying statute in objecting the EPA’s proposed action. These cases force EPA to think about how its interpretation of the statutory language will fare in court, making the law’s determinacy most important.

Second, the agency and its commenters are undoubtedly thinking about how interpretations will fare in court at every stage in the process, even before the agency ever initiates rulemaking. This greatly complicates our ability to apply game-theoretic intuitions about how sequentially rational, backward-induction-observing players in this game should operate—and perhaps makes some readers uneasy about what we can correctly infer from observing actions at just one stage in this larger interaction.¹⁴ I recognize these concerns—but see no reason why they should call into question the many cases in which we see legal constraint at work, even given the possibility that constraint could have happened invisibly at some other stage.

And, in fact, we do see a process of legal constraint observing simply by paying attention to the comments, EPA’s responses to them, and any subsequent litigation that seeks to have courts make legal claims effective constraints. Rulemaking routinely involves contestation. Figure 1 shows how environmental and industry groups differ in their propensity to comment on proposed rules, raise legal objections as part of their comments, achieve direct influence through those comments, and initiate litigation after promulgation of the rule. The EPA received adverse comments from environmental commenters in 38% and from industry commenters in 54% of

¹⁴ For an attempt to work through these issues, see Jordan and Rothenberg (2011).
rulemakings (Figure 1a), while the figures for legal objections are 26% and 33%, respectively (Figure 1b). (Where commenter identity is difficult to discern, the labels “environmental” and “industry” are used in a somewhat stylized manner, as meaning “for more stringent regulation” and “for less stringent regulation,” respectively.)

**Figure 1 - Contestation in Notice and Comment Rulemaking for CAA rules with nationwide applicability, 1991-2009**

- **a) Any Adverse Comments**
- **b) Legal Objections**
- **c) Legal Concessions, given Legal Objections**
- **d) Litigation, given Legal Objections**
- **e) Judicial Opinions with some Victory for Petitioner**
- **f) Litigation Producing Settlements**

Note: In a-d, circles represent estimated means and dotted lines represent 95% confidence intervals. For e-f, circles show true proportions (since all relevant cases are included in sample).
Most proposed rules do not receive legal objections, but many rulemakings under the CAA (even those of nationwide applicability) are of a fairly quotidian nature, and perhaps are simply not all that contentious. Not surprisingly, more important rules receive legal objections more often—using length as a serviceable proxy for importance, a rule of typical length has a 26% chance of receiving legal objections from environmental commenters and a 40% for industry commenters, while the corresponding figures for a rule twice as long are 36% and 85% (see Table 3, appended).

EPA mostly responds to legal objections with defenses of its original interpretations, suggesting that it is generally committed to the positions staked out in its proposals. It does make legal concessions in non-negligible portion of cases, though, possibly in order to avoid having its choices reversed in court. Given a legal objection by environmental commenters, EPA makes concessions in its legal interpretation in about 16% of rulemakings; the corresponding figure for industry is considerably higher, at around 38% (Figure 1c). In part, this discrepancy could be a response to industry commenters’ greater litigiousness. Given the presence of some legal objection, litigation is fairly likely—indeed, it is considered a normal part of the process, and need not signify any acrimony (Coglianese 1996). Environmental commenters bring litigation in approximately 30% of cases in which they raise legal objections, while industry commenters follow up their legal complaints with lawsuits about 37% of the time (see Figure 1d). If EPA’s fear of losing in court is a primary motivation for making legal concessions, the fact that industry litigants with greater resources can pursue their legal challenges more often should push the EPA to make pre-litigation concessions to industry more often.

When litigation does arise, it yields results for petitioners challenging an aspect of EPA’s interpretation often, although by no means as a matter of course. If a judicial opinion is reached,
environmental petitioners win on at least part of their claim in 63% of cases, while industry wins at a lower rate, 55% (see Figure 1e). Rather than pursuing disputes all the way through a costly trial, the EPA and its challengers often resolve their disputes through settlements. About 37% of cases brought by environmental commenters and 60% brought by industry commenters lead to settlements (see Figure 1f).\footnote{This is not quite the same as saying that these percentages of cases “settle” and therefore fail to produce a judicial opinion, because many challenges to rulemakings lead to both settlements and judicial opinions.} The agency would be unlikely to make concessions unless it felt there was some chance it would find itself on the losing end of a judicial opinion. (If the agency was simply inclined to make the concession on the grounds that doing so was good policy, it would almost certainly have done so in the finalization of the rule, thus avoiding costly litigation altogether.)

The claim that a healthy minority of CAA commenters is able to activate the statute’s legal constraints against the agency hardly seems to cry out for headlines, and yet it is an important finding. Learned observers (e.g., Posner and Vermeule 2010, chap. 3; DeMuth 2012) and loud-mouthed pundits alike frequently portray America’s administrative state as out of control, and subject to little meaningful control—perhaps especially for matters of low political salience. It is not uncommon to hear assertions of the nature: “Nobody wins legal challenges against the EPA,” and such claims implicate fundamental questions about our policymaking system’s accountability. One cannot respond convincingly to these claims by citing specific cases in which an agency was held accountable, since these can be dismissed as “exceptions that prove the rule.” The data presented here present a more serious challenge to these views, however, by showing that legal constraint can be regularly observed—even when we artificially confine our attention to just one stage in the policymaking process. More data from multiple
statutes and multiple agencies would be necessary to provide a global answer, but the evidence from the CAA rulemaking provides an important start.

**Politics as Usual?**

Even in light of the evidence just presented, those who are skeptical of law’s importance may at this point wonder whether all of this conflict is best understood as politics playing out in the bureaucratic and judicial arenas, rather than as real examples of legal constraint. Perhaps this is all just Democratic appointees catering to environmental commenters and Republican appointees doing the bidding of industry, both at the comment stage and in court. If this intuition were to explain the patterns of conflict, we could conclude that constraints on the EPA were mostly political rather than legal. In Appendix A, I show that this is not the case: politics plays a distinctly modest role, leaving us good reason to turn to statutory language as an explanatory factor of legal constraint.

**Determinacy and Conflict**

Perhaps the most convincing way to proceed at this point would be to move on to a presentation of some powerful statistical tests establishing strong relationship between middling values of textual determinacy and a high likelihood of legal conflict. This, of course, would require me to present some measure of determinacy that was simultaneously susceptible to such tests and capable of capturing the dynamics I have argued for earlier. Because I can find no measure that fulfills both of these requirements, I reject this approach.

Amongst those scholars studying laws-as-outputs, one approach with some popularity is to simply measure the length of the applicable statutory language and argue that longer statutes
These scholars assert that more complicated, specific, and exacting commands to the executive branch require more words. They acknowledge that not every marginal word, phrase, or section adds to a law’s restrictiveness but maintain that on average the correlation is a solid one, making length a useful proxy for constraint. I believe this argument is mistaken, because it fails to account for Congress’s goal of constrained flexibility. As explained above, extremely constraining laws can usually be written quite tersely, while creating delineated but useful flexibility for policymakers requires more verbiage. Shortness is as likely to signify clarity of purpose and the simplicity of the problem being addressed as it is to signify a choice to leave all of the important decisions to those carrying out the law. Even admitting the adequacy of length as a proxy for constraint for the sake of argument, however, attempting to use it as a way of understanding law-as-determinant quickly runs into an insuperable difficulty: which length should be measured? In the rulemaking context, the question can be translated into concrete terms by asking which sections of the law are directly implicated in the rulemaking, but meaningfully answering this question turns out to be nearly impossible. What will it mean for a section of the law to be directly involved? How will such involvement be detected? In the cases where legal disputes persist until a judicial opinion is issued, it may be easy enough to say that those sections that the court discusses as being directly relevant are involved. But even where this clarity is available, it fails to resolve the question of whether other sections of the law were equally relevant but less controversial. And then there is the question of what a “section” really consists of—especially when some formally designated sections of the law include tens of thousands of words and others just a few hundred. It is unclear whether these questions have answers, even in principle. This latter difficulty applies not just to length, but to nearly any type

---

16 In the game theoretic language of, e.g., Huber and Shipan (2002), they are said to represent smaller compliance intervals for the agent to whom power is delegated.
of mechanical content analysis, such as measuring the ratio of “shall” versus “may” or the relative prevalence of judgment-requiring adjectives. In short, because legal contestation is not rooted in undifferentiated blocks of statutory text, attempting to understand a law’s effects in this way is deeply problematic.

Without any satisfactory means of measuring determinacy mechanically, I proceed by choosing three sections of the CAA that are relatively coherent and self-contained, such that they clearly govern many rulemakings nearly by themselves. I parse each of these sections quite closely in order to get a sense of their determinacy, and argue that two exhibit medium determinacy—largely because of the use of adjectives in pivotal phrases—while the third combines high and low determinacy. I show how the medium-determinacy language gives rise to greater legal conflict, with the practical meaning of these sections largely defined through legal contestation.

Section 109 – National Ambient Air Quality Standards

One of the CAA’s central provisions is § 109, which requires the setting of National Ambient Air Quality Standards (NAAQS) for the most important pollutants. Achieving these standards is the condition for air quality “attainment,” which every part of the country must meet or attempt to meet through a set of costly measures specified elsewhere in the act. The act leaves no discretion to avoid the responsibility of issuing standards. The key phrase of the section is found in §109(b)(1), which says that the NAAQS must set standards, “the attainment and maintenance of which in the judgment of the Administrator, based on such criteria and allowing an adequate margin of safety, are requisite to protect the public health.” Although the requirement to set standards is clear, there is no specific definition of “public health” in the act,
leaving open the question of how strict a requirement this is meant to be. Is it “requisite” for the standard to guarantee perfect protection of health, such that any detriment must be eliminated, or does a merely “adequate” margin of safety imply that such absolutism is unnecessary? The latter interpretation is supported by contrasting this section’s “adequate margin of safety” requirement with the language of §112, discussed below, which requires standards securing an “ample margin of safety.” Since “adequate margin of safety” is not absolute, then, the phrase’s inclusion creates a clear example of medium determinacy: this is meant to be binding upon administrative discretion, but the precise nature of the constraint is unclear. On the other hand, EPA’s hand is strengthened by the inclusion of the phrase “in the judgment of the Administrator,” which suggests that deference to the agency’s findings is warranted.

The remainder of the section (which in all is less than 700 words) goes on to create some highly determinate requirements that EPA set and revise the NAAQS periodically, as well as creating a committee of scientific experts charged with reviewing the standards and issuing recommended revisions every five years if warranted.

A close reading of this relatively brief section shows that it combines very high determinacy in its requirement for setting standards with medium to low determinacy as to the actual content of those standards. The expectation would be for a fairly high level of contestation, although because of the section’s language favoring deference to agency choices, legal challenges could be expected to face an uphill battle.

Sections 112 and 129 – Maximum Achievable Control Technology

In addition to conventional air pollutants that create health risks through long-term exposure to outdoor air, from its adoption in 1970 the CAA also addressed the release of air
toxics, pollutants hazardous to human health even in trace quantities. Rather than setting permissible levels of these pollutants, the Act instructed EPA to adopt technology-based standards under §112 that would force all sources of air toxics to reduce their emissions as much as possible through adoption of the “maximum achievable control technology” (MACT).

Evidently, though, the 1970 version of the statute was insufficiently constraining in forcing regulation in the first place. Prior to the 1990 Amendments, EPA had promulgated just seven standards, and regulation of air toxics was widely regarded as a failure (Graham 1985). After the 1984 tragedy at a Union Carbide plant in Bhopal, India, in which an accidental release of methyl isocyanate killed thousands, a political consensus emerged to require more from the EPA (Bryner 1993).

The 1990 amendments were clearly aimed at constraining the agency’s discretion, which had in the past allowed it not to act. One of the amendments’ champions, Representative Henry Waxman (D – CA), penned an article (1991) that stated this in no uncertain terms. The new §112(b) listed 189 chemicals EPA would be required to regulate. Recognizing the ambiguity in “maximum achievable control technology,” the amendments also sought to offer a constraining definition of the concept in the new §112(d)(3), which required that the MACT “for new sources in a category or subcategory shall not be less stringent than the emission control that is achieved in practice by the best controlled similar source, as determined by the Administrator.” “Best controlled similar source” is further defined as meaning those sources whose average emission limitations are in the top 12% of their industrial category, or, where fewer than 30 sources exist, the average limitation achieved by the best five sources. A new §129, requiring EPA to set emission standards for solid waste combustion, copied this language nearly exactly.
Did these changes successfully deliver a determinate law that would give EPA clear, uncontestable directions? Several factors leave this section of the CAA well short of high determinacy. First, the basic definition of MACT in §112(d)(2) still requires standards that “the Administrator, taking into consideration the cost of achieving such emission reduction, and any non-air quality health and environmental impacts and energy requirements, determines [are] achievable.” This is fairly open-ended, although it is subject to the more objective standards of §112(d)(3). That section’s requirement that MACT be defined by reference to the “best controlled similar source” also relies rather heavily on the adjective “similar.” It is unclear what it means for sources to be “similar,” such that they will be grouped in the same category and thus subject to the same standards. Congress anticipated this (Waxman 1991, 1777), and thus in §112(c)(1) pushed the EPA toward using categories already broadly defined. That subsection’s language leaves some room for maneuvering, though:

To the extent practicable, the categories and subcategories listed under this subsection shall be consistent with the list of source categories established pursuant to section 111 and part C. Nothing in the preceding sentence limits the Administrator’s authority to establish subcategories under this section, as appropriate.

Here, the adjectives strike again: though the Administrator “shall” use the categories established pursuant to another part of the CAA, she must only do so when “practicable.” And, as per the final sentence, this “shall” is still further attenuated by insisting that the Administrator’s “authority” to define categories, “as appropriate,” should not fundamentally be called into question by anything in the section. Yet another source of ambiguity comes in defining what exactly constitutes a single source—is it a geographically proximate group of factories, just one factory, or just one self-contained unit within a factory? Answering this question has large implications for determining which industrial activities are grouped into categories together. Some of the flexibility afforded by these ambiguous words may have been consciously desired,
with legislators trying to give the EPA the ability to deal with a clearly complex and technical subject. It seems likely, though, that at least some of the flexibility resulted from the costs of trying to achieve even greater precision and the inherent slipperiness of language.

The 1990 Amendments (§112(f)) also require EPA to revisit each type of emissions source category after MACTs have been set and consider whether “beyond-the-floor” controls are required to “provide an ample margin of safety to protect public health.” As with §109, determining what constitutes an “ample margin of safety” is sure to be the source of legal conflicts.

In sum, §112 and §129 as amended in 1990 aimed to saddle EPA with some very highly determinate requirements that would force aggressive regulation. As the preceding discussion shows, though, many ambiguities remained in the statute, reducing the determinacy of key subsections.

Title VI – Protection of the Stratospheric Ozone

One of the major additions to the CAA in the 1990 Amendments was Title VI, which implements the terms of the Montreal Protocol, an international treaty created to protect the stratospheric ozone layer from substances that had caused its depletion, such as chlorofluorocarbons (CFCs). The title, which runs approximately 10,000 words, meticulously sets out definitions and specifies exactly which substances would be banned entirely and which would be phased out. It also provides an objective, determinate standard for adding other chemicals to the banned list, and a loose, discretionary standard for adding to the phase-out list “any other substance that the Administrator finds is known or may reasonably be anticipated to cause or contribute to harmful effects on the stratospheric ozone layer” (§602(b)). Citizens may
petition for addition or deletion from these lists, and the Administrator is compelled to respond with scientific justifications.

What adds to the title’s complexity is the need to create procedures for issuing exemption allowances for essential uses of the products, many of which lacked close chemical substitutes. The Administrator “may” authorize limited amounts of some banned chemicals “solely for use in essential applications…for which no safe and effective substitute is available” (§604(d)(1)). The use of “may” means that no legal challenge attempting to force EPA to issue such an exemption would stand a chance, while the rather vague adjectives used in delimiting its discretionary power to do so could lead to some questions as to whether the agency was acting within its authority. On the other hand, the next subsections create non-discretionary duties to issue exemptions for medical uses (mainly for the metered-dose inhalers used by asthmatics), in consultation with the Food and Drug Administration, and for airplane deicing, in consultation with the Federal Aviation Administration. These responsibilities to grant exemptions, along with some others, are constrained by another section that clearly states that “under no circumstances” are total exceptions to allow any producer to produce more than during a set baseline year (§604(d)(4)). The story is much the same for chemicals on the phase-out list, and for various other exempted uses. EPA is given some mandatory and some discretionary powers to grant production or consumption allowances, with specified criteria on when it can grant these exemptions and how large they can be in total. At the same time, its specific choices of which companies to grant allowances are fairly unconstrained.

Further parts of the Title create a number of other responsibilities for EPA, including: creating a trading exchange for allowances; establishing a program for recycling and disposing of appliances containing certain refrigerants; regulating the servicing and maintenance of car air
conditioners; and banning “non-essential” uses of all covered chemicals. EPA must also make it illegal to replace phased-out or banned substances with substitutes “which the Administrator determines may present adverse effects to human health or the environment, where the Administrator has identified an alternative” that (1) reduces overall risk, and (2) “is currently or potentially available” (§612(c)). While “non-essential” and “potentially available” certainly leave room for differing interpretations, most of these sections are relatively determinate in laying out the agency’s responsibilities and discretionary powers.

In short, Title VI is constructed in a way that tends to diminish legal contestation: it makes some things very clear, so that it would be hard to imagine the agency ducking these constraints. At the same time, where it builds in flexibility for the administrator, it does so with language that is unlikely to leave petitioners with much ammunition to challenge.

*Legal Conflict for §109, MACT, and Title VI*

It now remains to show how the determinacy of each of these sections of the CAA translates into legal conflicts and episodes of contestation.

The standard-setting of §109 clearly delineated a responsibility for the EPA to act, and EPA has never contested this duty. On the other hand, there has been a steady drumbeat of lawsuits to force the agency to conduct its reviews in a timely manner—showing that there is more to achieving acceptable policy results than agreeing on the presence of legal responsibility. On the other hand, the statute’s medium to low determinacy requirements for the standards left room for considerable legal conflict about the practical meaning of “adequate margin of safety.” Sure enough, since this part of the Act became effective in 1970, there have

---

17 Such lawsuits are not included in my data, because they do not arise from petitions for review of agency rulemakings and rarely can be said to involve any legal interpretation.
been 19 rulemakings directly setting NAAQS, and nearly every one featured serious legal objections raised by one side or the other, and usually both. Litigation followed in 11 (57.9%) of these cases (see Figure 2). This high level of legal contestation is undoubtedly at least partially because of the far-reaching implications of NAAQS; large potential benefits make legal contestation more attractive, regardless of the prospects of succeeding. Nevertheless, if the law were more determinate, we would expect a lower level of contestation.

The legal challenges have a mixed record of success at imposing meaningful constraints on the EPA. The D.C. Circuit has struggled with the ambiguity inherent in the pivotal language, trying to figure out just what sort of limit the requirement to secure an “adequate margin of safety” creates. Indeed, a divided D.C. Circuit panel in *American Trucking Associations v. EPA* (1999) felt that this section of the law was so lacking in determinacy, as interpreted by the EPA, that it ran afoul of the non-delegation doctrine of constitutional law. This dramatic holding was reversed by a unanimous Supreme Court in *Whitman v. American Trucking Associations* (2001), but the high court’s opinion hardly resolves the inherent indeterminacy. Rather, the majority opinion suggests that “Section 109(b)(1) of the CAA, which…we interpret as requiring the EPA to set air quality standards at the level that is ‘requisite’—that is, not lower or higher than is necessary—to protect the public health with an adequate margin of safety, fits comfortably within the scope of discretion permitted by our precedent” (475-76). But since neither “public health” nor “adequate margin of safety” ever receives a clear definition, this is hardly a model of clarity.

---

18 Because of the small sample size of §109 rulemakings in the large-n study, these data are based on a comprehensive search of NAAQS rulemakings since 1970, guided in part by the summary tables compiled by Bachmann (2007).
EPA standards were struck down by the courts in other parts of *American Trucking* that the Supreme Court left alone, and in *American Farm Bureau Federation v. EPA* (D.C. Cir. 2009), where EPA failed to convince the court it possessed evidence sufficient to show that its standard would “protect public health with an adequate margin of safety.” Similarly, EPA’s
Ozone NAAQS, proposed in 2007 and promulgated in 2008, was attacked by both environmentalists and industry interests as getting “adequate” wrong. Tied up in litigation, EPA decided to reconsider, eventually proposing to make the standard more stringent (EPA 2010), but it will probably face more legal challenges once it finalizes this standard.

On the other hand, none of these losses in court stands out as a rebuke that would force EPA to fundamentally change how it arrives at NAAQS decisions. Because the language is fairly ambiguous, courts are rarely able to say decisively that EPA has flouted the language and must reach a different conclusion. Instead, legal victories for petitioners simply force EPA to go back to the drawing board and propose its standards anew. This can lead to situations in which petitioners win a legal victory but have little in terms of policy to show for it. Environmental commenters exhorted EPA to strengthen its existing sulfur dioxide (SO$_2$) NAAQS, but the agency resisted and promulgated its decision to stand pat (EPA 1996). The D.C. Circuit found that the agency had failed to justify its choice not to create a standard for short-term bursts of pollution of special concern to asthmatics and others with impaired breathing function, and remanded the matter to EPA to address more thoroughly in *American Lung Association v. Browner* (1998). Action on remand was not immediately forthcoming, however, and so the petitioners were forced to return to the courts, where they settled with EPA in exchange for promises to regulate soon. This process then repeated itself—not once, but twice (EPA 1998, 1999, 2000). Not until 2006 did EPA finally solicit scientific input on the effects of short-term SO$_2$ exposure (EPA 2006). Another action-forcing lawsuit ended in a judicial decree that EPA propose a rule by late 2009 and finalize it by mid-2010 (EPA 2007). Finally, in a proposal coming nearly twelve years after the D.C. Circuit’s remand, the EPA proposed substantive revisions to the SO$_2$ standards (EPA 2009). In doing so, the agency finally addressed the need for
a 5-minute standard—and once again decided it was not appropriate, this time bolstering its conclusion with more scientific evidence. While the final rule did tighten standards significantly, environmental interest groups were disappointed on the matter of the 5-minute standard. This example vividly demonstrates how more than a victory in court is required to achieve effective legal constraint.

The MACT sections (112 and 129) were designed in the 1990 amendments to restrict EPA’s choices, but the analysis of the statutory language above shows that the attempt to provide constrained flexibility to the agency left plenty of ambiguities. These medium-determinacy parts of the statute could thus be expected to engender a great deal of legal contestation. The amendments were undoubtedly successful in prompting EPA to act: activities under these two sections account for 37% of all CAA rulemakings with nationwide applicability since 1990, in stark contrast to the two decades of inactivity prior to the amendments. Rates of legal contestation were just average, perhaps reflecting the fact that some rulemakings dealt with fairly marginal issues—overall, MACT rules were below average length. Given some legal conflict, though, rates of industry litigation were considerably higher for rulemakings under these sections. Industry commenters initiated litigation in 47% of cases in which they had legal objections (compared to 32% for all other cases).

The legal constraint EPA confronted in these cases fits well into the pattern described for medium determinacy statutory language. EPA was clearly and substantially constrained, losing a large number of cases dealing with MACT—60% of judicial votes cast in cases under these titles included some kind of industry victory, compared to 39% for all other parts of the law. These higher rates of success for legal petitioners undoubtedly encourage later petitioners to pursue their legal complaints. Perhaps because the agency has lost so many of these cases over the
years, it was also especially likely to settle industry-initiated cases dealing with §112 and §129. Once litigation was initiated, settlements followed a whopping 80% of industry petitions.

The many cases since 1990 have yet to definitively settle the ambiguities discussed above. Controversy has consistently surrounded attempts to define permissible criteria for determining “achievability,” with the only clear rule to emerge being that EPA may not nullify the MACT sections by setting a requirement of “no control,” even if the best performing sources up to that time used no technological controls specifically designed to target that particular type of emissions.19 The courts thus affirmed that EPA must choose some technological requirements for each kind of hazardous pollutant from each type of source, without offering much guidance as to how they should do so in the absence of past experience. The obligation to designate categories of “similar” sources for purposes of determining MACT has followed the same pattern: EPA’s attempts to determine subcategories have been litigated on numerous occasions, with a mix of wins and losses for the agency.

Finally, Title VI seemed to be the sort of statute that should lead to relatively little legal contestation, since most of its requirements seem either unambiguous or wholly discretionary. Indeed, we find legal objections registered by environmental commenters in just 14% of cases, well below average for the CAA as a whole. Of those rules that received legal objections in the comments, just 19% of environmental complaints and 11% of industry complaints led to litigation, both well below the average rates (see Figure 2). Those legal complaints, though, focused primarily on the few medium-determinate phrases noted above, including what exactly is covered by EPA’s duty to make “nonessential” products illegal and which chemicals can rightly be classified as “substitutes…currently or potentially available.”

To be clear, saying that there is relatively limited legal contestation under Title VI is not equivalent to saying that interest groups have especially poor chances of influencing policy set by EPA. Indeed, a whole host of issues about which the statute has little to say receive extensive comments, to which EPA is often receptive. Logistical questions about monitoring, recordkeeping, reporting, testing, verification, as well as questions about how the rule will deal with the state of the world before its existence, covering issues such as grandfathering, effective dates, and the relationship to other existing state and federal laws, may all be as important to those being regulated as questions of statutory interpretation. Regulated interests are likely to have the most specific information about technical issues relating to these questions, and so their input will be valued by the agency as it attempts to fashion a workable rule. Issues such as these can sometimes give rise to lawsuits that make little reference to the statute. For example, in *Arkema Inc. v. EPA* (D.C. Cir. 2010), a holder of tradable production allowances successfully challenged an EPA revision to the trading program as having retroactively altered the status of previously allowed transactions. Though this case produced one of the more vehement dissents out of all of the CAA challenges, the nature of the dispute was confined mainly to procedural rather than statutory requirements.

In those rare instances in which statutory interpretation challenges under Title VI were sustained against EPA all the way through to an opinion, EPA generally found itself on solid ground. Because the language in Title VI emphasizes the agency’s discretion to make determinations, legal adjudication ends up focusing on the adequacy of EPA’s scientific justifications for its decisions. Where it has good data (i.e., data that strikes judges as convincingly thorough), the agency could win lopsided victories, as in *OZ Technology Inc. v.*
EPA (D.C. Cir. 1997). If its justification was poor, it will suffer a remand on the specific question, but is unlikely to be bound by restrictive precedents constraining future interpretations.

The preceding discussion provides at least preliminary empirical support for the idea that medium determinacy statutory language is likely to engender the greatest amount of legal conflict. Sections relying on medium-determinacy language (§109, 112, 129) became the object of repeated legal contestation, while Title VI, constructed of a mix of high and low determinacy language, gave rise to far less legal conflict, especially in terms of litigation. Table 4 summarizes these findings.

<table>
<thead>
<tr>
<th>Section of CAA</th>
<th>Determinacy</th>
<th>Observed conflict</th>
<th>Litigation Rate (given legal objection)</th>
<th>Other Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>§109</td>
<td>Medium</td>
<td>High</td>
<td>55.6%</td>
<td>Back and forth record in court; Interpretation of “adequate margin of safety” remains uncertain</td>
</tr>
<tr>
<td>§112, 129</td>
<td>Medium</td>
<td>High</td>
<td>41.3%</td>
<td>Very high rate of settlements with industry challengers of what controls are actually “achievable”; High rate of litigation success for challengers</td>
</tr>
<tr>
<td>Title VI</td>
<td>Mix of low, high</td>
<td>Low</td>
<td>17.4%</td>
<td>Low rate of legal objections by environmental commenters; What few lawsuits there are tend not to be about statutory interpretation</td>
</tr>
</tbody>
</table>

To be sure, this evidence is illustrative rather than dispositive; proving the robustness of the relationship between determinacy and legal conflict to a skeptic would require an entirely different research strategy capable of systematically measuring and analyzing far larger amounts of data.

Some reassurance can nevertheless be found by consulting the wisdom passed down among specialized professional legislative drafters. The Legislative Drafter’s Desk Reference
warns that though adjectives and adverbs are often necessary, and sometimes politically demanded,

they can cause endless legal difficulties when used carelessly, because a modifier typically ascribes to a noun or verb a characteristic that it does not precisely define. … Other things equal, the fewer the modifiers the easier the statute is to administer. Every modifier that you use in a bill is likely at some stage to require an administrative or judicial interpretation (which may or may not accurately reflect the true legislative intent), no matter how aptly it is used or how much it smooths out the language. (Filson 1992, 254-55)

Conclusions

To understand which laws work best, we must see how well they perform their intended functions. Positing the function of regulatory laws as creating constrained flexibility for the executive branch, I have argued that understanding the effectiveness of constraints requires seeing what constraints can be effectively enforced against an unwilling interpreter through legal contestation. I further argue that legal contestation is likely to be most prevalent when laws feature language of medium determinacy—that is, language that clearly indicates that the law is meant to act as a limit on agency discretion, but is unclear about the precise nature of this limit. Medium determinacy is often the result of pivotal clauses’ reliance on judgment-requiring adjectives.

Given the relationship between medium-determinacy language and conflict, one might still think that the relationship between determinacy and constraint more broadly is intuitive: more determinacy leads to more constraint. Similarly, more determinacy means that Congress has solved more problems, while low determinacy laws leave the bureaucracy to flesh out the practical meaning of laws.20 Both of these intuitions are fundamentally sound, but they miss out on crucial insights by glossing over the role of legal conflict. First, they suggest that setting a

---

20 This closely tracks the insight that rules are ex ante efficient, while standards are ex post efficient (Parisi 2003). Much of this essay’s discussion tracks familiar arguments about bright-line rules versus flexible standards, but I have chosen to avoid these dichotomous abstract terms in favor of talking more directly about realized determinacy as a continuous concept.
The statute’s level of determinacy is a simple choice for the legislature. This frequently invites a follow-up question: why doesn’t Congress just write very highly determinate laws? The existing scholarship on delegation to bureaucrats gives some reasons based on a principal-agent framework. This chapter suggests, though, that Congress’s motivations are more complex and its abilities to prevent legal conflict are more limited than is generally imagined. If the aspects of the world legislators are hoping to change are complex enough that they cannot in good conscience (or in good politics) choose rigid rules unlikely to fairly or effectively balance their intended goals, then they end up trying to get some determinacy through words that admit flexibility. Their use of these words engenders litigated conflict over what constraints those words actually create.

Given these complexities, determining how one part of a statute will play out can be very difficult for legislators. Attempts to nail down details such that they create no room for differing interpretations often fail, leaving room for contestation. Conversely, broad mandates may sometimes empower (uncaptured) agencies to resist legal attacks, leaving them with the discretion needed to effectively tackle complex problems—though for those who would challenge their interpretation, this positive characterization of agency discretion will ring hollow.

There are also institutional consequences for legal determinacy. A workable compromise between legislatively-specified preferences and bureaucratic attempts to evade them may be forged through legal conflicts—but then, disputes may end up producing resolutions that few would have desired. Melnick (1983) gives many examples of how federal judges took the Clean Air Act in directions few people anticipated, sometimes with long-lasting consequences for the shape of national air pollution policy. Katzmann (1986) argues that ambiguous statutory language addressing the problem of transportation disability policy, which itself grew out of a

---

21 As Chapter Five shows, rigidly constructed rules are also likely to suffer from statutory erosion.
hazy definition of the nature of the issue, led to costly back-and-forth legal battles and ended up empowering government lawyers at the expense of policy experts.

More generally, judges are put into a difficult position when charged with adjudicating disputes over medium-determinate language. They can abdicate responsibility by simply accepting the agency’s professions of good faith in interpreting the law, in which case they risk allowing serious mischief by executives. Or, they can effectively legislate the specific meaning that Congress failed to provide, lowering the perceived legitimacy of policies and opening the door to charges of “judicial activism.” On the other hand, if the judiciary is cut out of the conflict-resolution loop through settlements, this presents serious issues of institutional legitimacy in itself, as it will seem that policy becomes a function of who has the most willingness to credibly threaten to pursue costly litigation rather than who can prevail in transparent, merit-based public comments (Rossi 2002; Fisher and Schmidt 2001).

This portrait of legal constraint in action looks very different from the one painted by “congressional dominance” scholars, who portrayed Congress as able to control the path of future legal development through a careful mix of precise statutory language and required administrative procedures. That picture makes legislators seem like wizards—or really, like a single, unified wizard. Like Obi-Wan Kenobi, this wizard looks into the future and sees he won’t be there, but devises ways to ensure that even in his absence everything will turn out as he foresees. Legislators are not so strong in the force, though—neither so clever nor so clear-minded as Obi-Wan. The choice of determinacy is often dictated by necessity rather than a desire to confer regulatory flexibility; and whether flexibility is intentionally included or manages to sneak in, it is likely to be put to unanticipated uses. We should be wary of imagining legislators appearing in spirit after every court decision to beneficently applaud the realization of
their own plans. In any case, it isn’t clear that the political rewards for legislators demonstrating extraordinary foresight in their drafting choices are very great. At the risk of badly mixing metaphors, former HEW Secretary Joseph Califano expressed a similar point nicely, saying of legislators, “It was more fun [for them] to be Moses and deliver the commandments than to be the rabbis and priests who had to make them work” (quoted in Katzmann 1986, 57-58).

Keeping this more realistic picture of Congress in mind, are there recommendations for lawmakers that flow directly out of this chapter’s analysis of determinacy and conflict? There are a few, but they are unlikely to make for exciting reform proposals or bumper-sticker-ready political messages. Congress should be clear-eyed about the consequences of choosing medium-determinacy language as a mechanism of control. It is easy to become enamored of the idea of supple “standards” instead of brittle “rules,” but often language that makes clear the existence of an obligation without specifying the nature of that obligation with any precision leads to drawn-out legal disputes to determine what the effects of the law will be. If Congress wants courts to take the lead, this is certainly a viable option; ambiguous phrases can be worked into determinate terms of art through a judicial common law process, and there may be policy areas in which entrusting generalist judges with this function is a sound choice. For many complex technical issues in modern regulatory policy, however, there are reasons to think that specialized bureaucrats may be better equipped to develop policies under low-determinacy laws. Sunstein (1998) suggests, approvingly, that in America’s modern administrative state, executive branch agencies have become the primary developers of the common law.

Of course, while the flexibility low-determinacy laws give to bureaucrats can be a virtue in forging a coherent policy in a complex policy area, it can also be a vice. As Chapter Four shows, it may open the door for laws to be used for purposes legislators never imagined and
perhaps would not approve. On the other hand, as Chapter Five shows, it may allow bureaucrats to abandon the legislation’s goals while remaining obedient to its letter. All this is very well captured by standard political science principal-agent models of bureaucratic discretion (e.g. Huber and Shipan 2002). What this chapter’s discussion adds to these models is a sense of how legislators should proceed if they hope to clearly delimit a zone of discretion for bureaucrats through legal constraints—that is, if they believe achieving control in a principal-agent setting is an important goal. Basically, legislators must focus their attention on highly-determinate language that will “hard-wire” certain features of the regulatory regime. Law as constraint is utilized to its greatest effect when it is used to prevent mission creep, in either direction (toward imperialist ambition or neutralization through capture). But where such detailed choices will be necessary so as to make it impossible to write determinate rules at the front end of the legislative process, there is little reason to think that medium-determinate language offers some sort of golden mean that will direct and constrain bureaucratic actions, but not too much. Such language is unlikely to provide certain guidance and will therefore lead to confusion and disagreement, and thence to awkward muddling through and legal conflict over the statute’s “true” meaning.

Whatever the worth of the preceding paragraphs, they seem unlikely to change legislators’ thinking about how to legislate. The incentives for using phrases like “adequate to protect public health” outweigh the remote-sounding risks of legal confusion warned of here. As with this dissertation’s other chapters, though, I end by considering whether there might be recommendations of more marginal changes that policymakers could be persuaded to adopt, if presented with the right evidence.
It would be ideal if a much larger set of statutory provisions could be investigated in an analogous manner, lining up the determinacy of key phrases with the section’s tendency to generate legal conflict. If this base of knowledge were broad enough, it might yield generalizable insights that the current work only hints at. For example, there might well be regularities in how determinate various medium-determinate words are. Perhaps “necessary” is a word that tends to induce discretion to bureaucratic judgments, while “reasonable” makes judges more likely to vindicate challenges to agency interpretations. One might suppose that objective sounding words such as “achievable” would be more contestable, but whether this intuition is correct remains an open empirical question. A global investigation of what sort of legal treatment different language receives would reveal patterns that could greatly aid drafters looking to lock in more or less determinacy.

A larger base of knowledge would also afford the opportunity to test whether ambiguities are resolved over time. Even if a law’s phrases fail to constrain legal interpretations when they are drafted, we might expect that receiving several judicial opinions all construing a particular phrase similarly might provide determinacy going forward. If so, litigation contesting the fundamental meaning of statutory language should be front-loaded and diminishing over time. Rulemaking since the 1990 CAA Amendments fits this theory: the percentage of rules litigated decreases over time, even controlling for rule importance (as proxied by length). In other cases, though, parties may be emboldened to test the constraints of a statute as the political coalition that passed the law recedes into memory, in which case there might at some point be a resurgence of litigation (as is the case in Chapter Five). Understanding the connection between particular statutory language and the patterns of conflict generated over the lifetime of a statute would allow drafters to avoid historically troublesome formulations.
More generally, this chapter demonstrates the potential productivity of further inquiry into statutes-as-determinants, rather than statutes-as-outputs. A program of research in this vein has great potential to aid policymakers charged with drafting laws by helping them to understand the likely consequences of their choices once they are subjected to legal and political contestation.
Table 3 – More Significant Rules (proxied by length) Receive More Legal Objections

<table>
<thead>
<tr>
<th></th>
<th>Environmental Legal Objections</th>
<th>Industry Legal Objections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of Final Rule in</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Westlaw Screens</td>
<td>0.193** (0.067)</td>
<td>0.839** (0.215)</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.540** (0.281)</td>
<td>-2.484** (0.449)</td>
</tr>
</tbody>
</table>

Logit coefficients reported, with standard errors in parentheses. * p < 0.10; ** p < 0.05

Predicted Probabilities for Rules of Various Lengths.

<table>
<thead>
<tr>
<th>Length</th>
<th>Environmental Legal Objections</th>
<th>Industry Legal Objections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short (1 screen)</td>
<td>20.6%</td>
<td>16.2%</td>
</tr>
<tr>
<td>Mean (2.47 screens)</td>
<td>25.7%</td>
<td>39.8%</td>
</tr>
<tr>
<td>Long (5 screens)</td>
<td>36.0%</td>
<td>84.7%</td>
</tr>
</tbody>
</table>

Appendix A – Politics as Usual?

The intuition that partisanship drives variation in agency responses to legal contestation finds support in the data—but only selectively and weakly. It seems that the affinity between Democratic administrations and environmental commenters leads these parties to more frequent agreement even before the notice-and-comment rulemaking process. Environmental commenters made legal objections in just 19% of Clinton administration proposed rules compared to 34% of Bush 43 administration rules; symmetrically, industry was also more likely to raise legal objections against Clinton administration proposed rules (35%) than Bush 43 administration proposals (23%) (see Table 5a). Given that they do register legal concerns, though, each side’s ability to influence bureaucrats was roughly equal (Table 5b). Apparently, for the most part, once the EPA and interest groups have parted ways on legal interpretations at the proposed rule

---

22 Curiously, while the data show that while there are no discernible differences between the Clinton and Bush 43 EPAs, environmental commenters were more likely to be influential in the Bush 41 administration than either of the following two, while industry commenters were more likely to be influential in both the Clinton and Bush 43 EPAs compared to the Bush 41 and Obama EPAs (which accounts for the statistical significant findings in Table 4b). Because of the relatively small amounts of data collected from the Bush 41 and Obama administrations, though, I am hesitant to base any substantive conclusions on these patterns.
stage, articulation of those disagreements is unlikely to be especially influential even coming from favored constituencies. When litigation did occur, the Clinton administration was

Table 5 – Legal Objections and EPA Responses by Administration
Maximum Likelihood Estimations of logit models of interest group legal objections by group type and proposing administration, CAA Rulemakings with Nationwide applicability, 1991-2009

- a) There are differences in patterns of legal objections

<table>
<thead>
<tr>
<th>Administration Proposing Rule:</th>
<th>Environmental Legal Objections</th>
<th>Industry Legal Objections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinton</td>
<td>0.329 (0.793)</td>
<td>-0.807 (0.643)</td>
</tr>
<tr>
<td>Bush 43</td>
<td>1.165 (0.782)</td>
<td>-1.492**(0.636)</td>
</tr>
<tr>
<td>Length (in Westlaw screens)</td>
<td>0.226** (0.069)</td>
<td>0.798** (0.211)</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.315** (0.787)</td>
<td>-1.433* (0.740)</td>
</tr>
</tbody>
</table>

Logit coefficients reported, with standard errors in parentheses. * p < 0.10; ** p < 0.05

- Predicted Probabilities by Proposing Administration

<table>
<thead>
<tr>
<th>Proposing Administration</th>
<th>Environmental Legal Objections</th>
<th>Industry Legal Objections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinton</td>
<td>19.4%</td>
<td>34.6%</td>
</tr>
<tr>
<td>Bush 43</td>
<td>34.3%</td>
<td>23.1%</td>
</tr>
</tbody>
</table>

- b) Given the presence of legal objections, there are not significant differences in whether they lead to concessions (see footnote 27)

<table>
<thead>
<tr>
<th>Administration Finalizing Rule:</th>
<th>Environmental Legal Objections</th>
<th>Industry Legal Objections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinton</td>
<td>0.189 (0.774)</td>
<td>1.627**(0.648)</td>
</tr>
<tr>
<td>Bush 43</td>
<td>-0.178 (0.745)</td>
<td>1.631**(0.667)</td>
</tr>
<tr>
<td>Length (in Westlaw screens)</td>
<td>0.105* (0.059)</td>
<td>0.047 (0.078)</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.137** (0.700)</td>
<td>-2.053** (0.632)</td>
</tr>
</tbody>
</table>

Logit coefficients reported, with standard errors in parentheses. * p < 0.10; ** p < 0.05
significantly more likely to settle with environmental challengers compared with all other administrations, but no other statistically significant patterns are present in settlements (data not shown).

Might it be that EPA-as-CAA-rule-maker is resilient to the pressures of political appointees, leading it to steer a middle course, but politically-appointed judges exert recognizable partisan influences? Certainly, a significant body of literature empirically assessing the record of government agencies in appellate courts has suggested that political patterns are recognizable and significant. If political dynamics were the main drivers of judicial review of agency decisions, we would expect that the agency would attempt to anticipate this behavior by making more concessions to industry petitioners when the D.C. Circuit was nearly full of Republican appointees (as it was in the early 1990s and toward the end of the Bush 43 administration) compared with times when the court was more balanced in its partisan composition (toward the end of the Clinton administration). During such times, petitioners would be more likely to face three-judge panels with two or three Republican appointees. Conversely, we would expect to see that a higher proportion of Democratic appointees would correspond to a higher rate of concessions to environmental commenters’ legal objections. There is no support for these hypotheses in the data. A simple logit model testing only the effect of the composition of the D.C. Circuit during the year of final rule promulgation finds that more Democrats on the court leads to greater influence (both legal and total) for industry commenters,

\[\text{This literature follows the example of Schuck and Elliott (1990), who constructed a wide dataset of appellate court opinions in hope of understanding whether the case of } \textit{Chevron v. Natural Resources Defense Council} (1984) \text{ had effectively changed judicial behavior. Revesz (1997, 2001), Stephenson (2004), Miles and Sunstein (2006), and Sunstein and Miles (2008) all use related methodologies to argue that the political ideology of judges plays a large role in their decisions. O’Leary (1993) takes a broad look at how the EPA fares as a litigant in the federal courts. Because these studies encompass only those cases litigated all the way to published opinions, however, they are limited in their ability to support broader inferences about the effects of the law on agency policymaking.}\]
the opposite of what was hypothesized, and this finding disappears when controlling for the administration which finalized the rule (see Table 6).  

There is somewhat better evidence for political dynamics in the judiciary when we turn our attention directly to judges, but the effects are still quite small. For each case that produced a judicial opinion, analysis is conducted at the level of each judge’s vote (since this is the level at which choices are made). Each judicial vote is coded as a total vindication of EPA’s interpretation, a total repudiation, or mixed. Republican-appointed judges were somewhat more likely to fully vindicate the EPA against environmental challenges than Democratic appointees; similarly, Democratic appointees were somewhat more likely to completely reject industry petitioners’ arguments compared to their Republican-appointed counterparts (see Figure 3). But these effects are fairly subtle—certainly not a vindication of the idea that judges simply vote their appointing party’s preferences. There are also few dissents, and those that do exist again fail to tell a tale of extreme partisanship; panels with three Republican-appointed judges had more frequent full dissents than panels of mixed partisanship (see Table 7). Environmental petitioners fare better than their industry counterparts in front of both Democrats and Republicans, making it hard to believe that politics can hardly be the major driver of judicial opinions. In short, the evidence strongly suggests that, while it is naïve to ascribe partisan politics no role in determining the dynamics of legal conflicts under the CAA, this role is a distinctly limited one.

---

24 This counter-intuitive significant finding also disappears if I run the logit regression without weighting.
Table 6 – Did EPA’s concessions to legal objections strategically anticipate politicized rulings by the D.C. Circuit?

Maximum Likelihood Estimations of logit models of EPA concessions to interest group legal objections as a function of D.C. Circuit political composition, CAA Rulemakings with Nationwide applicability, 1991-2009

Did EPA make *more* concessions to environmentalist legal objections when there were *more* Democrats on the D.C. Circuit?

<table>
<thead>
<tr>
<th>Logit Model</th>
<th>1 – simple</th>
<th>2 – controlling for administration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of Democratic appointees on D.C. Circuit when rule finalized</td>
<td>3.206 (3.635)</td>
<td>3.124 (4.107)</td>
</tr>
<tr>
<td>Administration finalizing rule:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bush 41</td>
<td>-</td>
<td>2.518* (1.457)</td>
</tr>
<tr>
<td>Clinton</td>
<td>-</td>
<td>0.232 (0.967)</td>
</tr>
<tr>
<td>Bush 43</td>
<td>-</td>
<td>-0.085 (0.873)</td>
</tr>
<tr>
<td>Obama (omitted)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.889** (1.373)</td>
<td>-2.949** (1.437)</td>
</tr>
</tbody>
</table>

Did EPA make *fewer* concessions to industry legal objections when there were *more* Democrats on the D.C. Circuit?

<table>
<thead>
<tr>
<th>Logit Model</th>
<th>1 – simple</th>
<th>2 – controlling for administration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of Democratic appointees on D.C. Circuit when rule finalized</td>
<td>6.062* (3.525)</td>
<td>2.263 (4.344)</td>
</tr>
<tr>
<td>Administration finalizing rule:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bush 41</td>
<td>-</td>
<td>0.298 (1.134)</td>
</tr>
<tr>
<td>Clinton</td>
<td>-</td>
<td>1.610* (0.962)</td>
</tr>
<tr>
<td>Bush 43</td>
<td>-</td>
<td>1.637* (0.906)</td>
</tr>
<tr>
<td>Obama (omitted)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.786** (1.404)</td>
<td>-2.713* (1.486)</td>
</tr>
</tbody>
</table>

Logit coefficients reported, with standard errors in parentheses. * p < 0.10; ** p < 0.05
Table 7 – Judicial Dissents on 3-judge panels of D.C. Circuit in litigation stemming from nationwide CAA rulemakings petitions for review, 1991-2009

Panel Composition, by appointing president’s party

<table>
<thead>
<tr>
<th></th>
<th>3 D, 0 R</th>
<th>2 D, 1 R</th>
<th>1 D, 2 R</th>
<th>0 D, 3 R</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>With full dissents</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Any kind</td>
<td>0 0%</td>
<td>2 6.9%</td>
<td>2 4.0%</td>
<td>2 10.0%</td>
<td>6 5.9%</td>
</tr>
<tr>
<td>- Along predicted partisan lines</td>
<td>* *</td>
<td>2 10.3%</td>
<td>0 0%</td>
<td>* *</td>
<td>2 2.5%</td>
</tr>
<tr>
<td>With partial dissents</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Any kind</td>
<td>0 0%</td>
<td>3 10.3%</td>
<td>8 16.0%</td>
<td>0 0%</td>
<td>11 10.8%</td>
</tr>
<tr>
<td>- Along predicted partisan lines</td>
<td>* *</td>
<td>1 3.4%</td>
<td>7 14.0%</td>
<td>* *</td>
<td>8 10.1%</td>
</tr>
<tr>
<td>Total panels (n)</td>
<td>3 29 50 20</td>
<td>102</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Since all judges on these panels were appointed by the same party, there can be no dissents along partisan lines. The total percentages in the right-hand column are calculated accordingly.
Methodological Appendix

Unit of Observation: The unit of observation is a “rulemaking chain”—that is, a set of connected actions including a proposed rule, comments and legal objections on it, a final rule in which comments have been answered, and any subsequent litigation stemming from a petition for review of the rule, which may include a settlement and/or a judicial opinion (published or, rarely, unpublished). By focusing on all rulemaking activity under the statute rather than simply examining those rules that eventually face legal challenges, I avoid selecting on the dependent variable. I am thus able to support different kinds of inferences than the many existing studies that use only data from cases that reach final opinions (see footnote 23). Few other studies use such an expansive lens to understand the effects of contestation in the rulemaking process—though Magat, Krupnick, and Harrington (1986) conduct a study of the efficacy of comments in CAA rulemakings worth noting (but completely different in its focus than the present work).

Universe of cases: In order to facilitate easy comparisons, I use only rulemakings under the CAA with nationwide applicability from 1991 to 2009. This excludes much significant activity under the law, including guidance documents, orders, and enforcement decisions. It also excludes the majority (numerically) of rulemaking under the CAA, which consists of EPA review of submitted revisions to a state or air pollution control district’s state implementation plan under §110 of the CAA. In these proceedings, the primary interpreter is the state or local air pollution authority, and the EPA’s role consists of auditing those choices to ensure that they are

25 I searched the Federal Register with the following Westlaw query for each year (changing the date as applicable): ca.pr(epa “environmental protection agency e.p.a. & “proposed rule”) & air & da(2000). This search is quite over-inclusive, as many rulemakings not under the CAA, and all statewide rulemakings will be included. I selected those cases that actually belonged in the dataset through examination. Once I had a proposed rule, I could generally find the final rule by searching for the citation of the proposed rule, but often I had to improvise searches to find the final rule due to typos or irregularities in citation practices. I then found settlement notices by searching the Federal Register for citations to the final rule, though again sometimes settlements were turned up through more idiosyncratic means. Finally, I found judicial opinions by searching the opinions of the D.C. Circuit for citations to the applicable final rule, making sure that the cases found in this manner actually were petitions for review.
compliant with the underlying requirements of the statute and national rules. Consequently, this paper’s discussion should not be understood as an attempt to provide a comprehensive account of how the CAA influences government actions. Limiting the universe as I have done is instead meant to provide good material for the discussion of determinacy and legal constraint. Choosing only nationwide rulemakings limits analysis to a single procedure and judicial venue (because under §307(d) of the CAA petitions for review of nationwide rules go directly to the D.C. Circuit Court of Appeals). I further limit my analysis to those rulemakings with proposals from the years 1991-2009, so as to deal with a (nearly) unchanging statute. I include in nationwide rulemakings those that deal with large parts of the country, such as the Ozone Transport Region (which includes most of the eastern United States.)

Sampling: There are 810 rulemakings that fit into the universe just described. I conduct my analysis on a sample of 331 of these, including all 146 which led to any kind of litigation (18% of all rulemakings) and 185 of those that did not lead to rulemaking (27.8% of the rulemakings without litigation). This sample is composed of all known rulemakings giving rise to litigation, all of the rulemakings of any kind for four years (1993, 1998, 2001, and 2006)\(^{26}\), and a random sample of 10 percent of the remaining rulemakings from each of the other fifteen years. This procedure imitates the rare event analysis technique often utilized by political scientists, which allows scholars to expend their limited resources on carefully coding many variables for each case rather than spending a disproportionate amount of time coding low-

\(^{26}\) These four years were selected because of the variation they offer in terms of political contexts. While keeping constant the year of Presidency (1\(^{st}\) and 6\(^{th}\) for Clinton and Bush 43), they offer variation in divided government (unified in 1993, 2006, divided in 1998, divided in 2001 after Jeffords’ party switch), composition of the D.C. Circuit (R-heavy in 1993, 2006, more balanced in 1998, 2001), and party control of the EPA.
information cases (King and Zeng 2001). For all analysis, each rulemaking is given its appropriate sample weight.27

Variables: Each rulemaking chain is coded for the following information: date and citation of proposed rule and final rule, and corresponding administrations; subject heading of rule; cited sections of the CAA; governing title of the CAA (with §112 and §129 split out from other Title I, and Title V grouped with Title I), length of final rule measured in Westlaw screens; whether there were environmental and industry comments28, whether those included legal objections, whether the legal objections had any influence in the final rulemaking, whether the other comments had any impact in the final rulemaking, and what the comments were disputing; whether there was litigation, environmental or industry, and whether that litigation produced a settlement or judicial opinion (or both); whether any follow-up actions were taken directly as a part of this rulemaking chain; if there was a settlement, its date, citation, and substance. If there was a judicial opinion, somewhat more information was recorded about it, and then that data was transformed so that it could be analyzed by each judge’s vote rather than by case. Information included the case name and citation; the identity of the petitioner(s); the identity of each judge on the panel and the party of the appointing president; the identity of the opinion writer(s); whether there was a dissent; whether each judge ruled completely for EPA, completely for the petitioner, or a mix; what the remedy was, if relevant; which statutory provisions the court’s discussion

27 I use Stata’s pweight function for all of my estimates, weighting each observation with litigation or from the four selected years as 1, and weighing each other observation by the number required to simulate the total number of rulemakings from its year. For example, in 2000 there were 43 total rulemakings, 7 of which gave rise to litigation. Of the remaining 36 rulemakings, my data includes 4 of them, each weighted as 9. For my descriptive data, this is clearly the proper choice. For the logistic regressions, one could argue that I would need a different weighting procedure because I oversampled on the dependent variable; but not much of substance seems to ride on my choice of weights. When I run any of my logit regressions without weighting, my results are virtually unchanged.

28 As noted in the text, when commenter identity is difficult to discern with certainty, the categorizations of environmental and industry are used in a stylized manner—with comments pushing for more stringent regulation being coded as environmental and comments pushing for less stringent regulation being coded as industry. State commenters are coded as industry or, more often, as environmental if their comments clearly push in one of these directions, although in some cases their interests cannot be so neatly included in either group and are therefore omitted from the main analysis.
highlighted; whether standing, ripeness, or justiciability was raised as an issue, and whether any of these factors was used to disqualify some argument; whether *Chevron v. NRDC* was raised as a controlling precedent, and if so whether the agency won or lost at *Chevron* Step One or Step Two; whether the agency’s interpretations were questioned as arbitrary and capricious, and how this inquiry was resolved.

**Importance:** I recorded length, measured in the number of Westlaw screens taken up by the final rule, as a way to proxy the substantive importance of the rulemaking. This is far from perfect, as the variable almost certainly rises when adverse comments are plentiful, and thus fails to be completely independent from controversy. While the measure is rough, it displays a significant amount of variation, and it is quite clear that EPA provides longer preambles and longer regulatory text when dealing with especially important rules.

**Missing data:** For many rulemakings, ascertaining the true value of some variables proved difficult. For example, it was often difficult to determine whether there were any environmental commenters. I treat the absence of evidence as evidence of absence, although this undoubtedly must be wrong in some small portion of rulemakings. For the analyses conducted above, then, I assign a value of 0 or 1 for every applicable variable for every rulemaking. I do not extend this inference to those cases in which I was unable to locate a final rule resulting from the proposed rule, however. I am fairly confident that these proposed rules simply were not finalized, making questions of comments received very difficult to resolve (and questions of litigation moot). As a result, these few cases (just 10 of them) are omitted from quantitative analyses. Finally, there is some issue of having results time-censored, since some rules proposed in more recent years may not yet have had sufficient time to produce litigation. I largely counteracted this difficulty through customized searches in which I scoured all manner of
sources to determine whether litigation was pending for recent rulemakings, such as the Obama EPA’s various greenhouse gas rules. It is likely that a few cases will yet produce unexpected litigation, but there is no reason to think that this will alter any of my substantive conclusions.
Cases Cited


OZ Technology Inc. v. EPA, 129 F.3d 631 (D.C. Cir. 1997)


American Trucking Associations v. EPA, 175 F.3d 1027 (D.C. Cir. 1999)

Appalachian Power Co. v. EPA, 208 F.3d 1015 (D.C. Cir. 2000)


Center for Biological Diversity v. Johnson, Civ. No. 05-1814 (D.D.C. 2007)

Natural Resources Defense Council v. EPA, 489 F.3d 1364 (D.C. Cir. 2007)

American Farm Bureau Federation v. EPA, 559 F.3d 512 (D.C. Cir. 2009)

Arkema Inc. v. EPA, 618 F.3d 1 (D.C. Cir. 2010)

Regulatory Materials Cited

Environmental Protection Agency (EPA)


Works Cited


Chapter Four –
When Can You Teach an Old Law New Tricks?

“A law…never produces exactly the results that anyone would have desired. It falls short, overshoots, or goes clean off in some other direction.”
—William Letwin, Law and Economic Policy in America (1965, vi)

If you can’t teach an old dog new tricks, how about old laws? Statues, unlike canines or constitutions, can be straightforwardly amended by legislatures. But amendments have all the same procedural requirements as new laws, and so in a sense every amended law is new again. In this chapter, I consider attempts to turn unamended laws to new ends through changes in legal interpretation, thus changing policy without changing law. In doing so, I draw attention to an underappreciated dynamic of policy change; provide insight into who is likely to spearhead such efforts and when courts are likely to be sympathetic to the novel interpretations; and provide a coherent explanation reconciling two of the Supreme Court’s most consequential statutory interpretation cases in recent years, FDA v. Brown & Williamson Tobacco Corp. (2000) and Massachusetts v. EPA (2007), which has proven to be a difficult feat.¹

Studying the dynamics of “old law, new trick” illuminates how relative institutional capacities determine the demands on, and limits of, legal interpretation. In some idealized versions of our constitutional policymaking system (both normative accounts and simplified theoretical models²), legal interpretation is conceived of as a limited endeavor confined to working out details of implementation and fitting life’s endless complexities into existing legal

¹ For example, Moncrieff (2008, 595) argues, “There is…no coherent story about the legal and political circumstances underlying Massachusetts and Brown & Williamson that would reconcile the two holdings.”
² This dissertation’s Introduction catalogues a variety of models of law that view law almost entirely in terms of a principal-agent model in which the legislature (including the president’s involvement) is the principal and the executive and judiciary are its agents in implementation and interpretation. While I also see legislative supremacy as a central value in the American political system, I emphasize the limitations of thinking in principal-agent terms for understanding the practical effects of statutes.
categories. Momentous questions of policy substance, though, are the domain of the legislature, the body most directly accountable to the public and most able to balance competing priorities. Reality, however, is not so neat. Congress resembles many things, but an ideal of a unified lawmaker willing and able to address every policy problem in a timely manner is not among them. Where Congress is unwilling or unable to effect policy changes through new legislation, other actors seek alternative routes of securing change, one of which is to offer novel legal interpretations of existing laws. This technique offers an attractive means of adaptation, but taken to its logical extreme it has the potential to eviscerate the idea that laws not only empower, but constrain, the government.

In many ways, evaluating “old law, new trick” interpretations is no different from evaluating an interpretation of a brand new statute. Judges will accept a construction of a statute when presented with convincing evidence that the statutory text, intent, and purpose support it. In the cases discussed in this chapter, however, the evidence is insufficient to dispel ambiguity about the statute’s requirements. As a general rule, aging several decades makes a statute less likely to dispositively resolve interpretive difficulties as they arise. Opponents of a new interpretation characterize it as a sharp turn that the law’s framers would never have countenanced with at least some plausibility. Text, purpose, and intent, and even the doctrine of *Chevron* (1984) deference may all be indeterminate⁴, leaving judges to rely on other factors in making their decision.

In such hard cases, I argue that judges use institutional factors to weigh “old law, new trick” interpretations. Judges will reject interpretations they see as likely to displace a healthy, functioning legislative process. When they believe institutional frictions and dysfunctions

---

⁴ One of the central ambiguities surrounding *Chevron* is when the doctrine should even be applied, the so-called “Step Zero” question (Sunstein 2006), which Section II addresses.
necessitate circumventing the normal lawmaking process, however, they will accept “old law, new trick” as a second-best, pragmatic substitute for statutory changes made by a directly accountable legislature. Judicial disagreement centers on just how much friction is tolerable—as well as how decisive the legislature must be to make a conscious choice of government inaction.

My argument leads me to criticize theories of statutory interpretation that make universal assumptions about institutional capacities. For example, Adrian Vermeule’s *Judging Under Uncertainty* (2006), justly recognized as one of the strongest statements of textualism, argues that judges should strictly adhere to textual requirements but then modestly defer to executive branch constructions where there are statutory ambiguities. He justifies this position by arguing that judges are likely to commit as many errors of commission as of omission if they allow themselves to think more ambitiously about the constructive role that creative interpretation might play; in other words, they would begin to “correct” problems rightly left to the legislature to address, if indeed they require addressing. I argue that practicing judges are unlikely to base their decisions on such sweeping judgments about institutional capacities. Instead, they will assess the institutional capacities relevant to the case at hand. When they believe that a legislature is unlikely to address the problem at hand, they are more likely to allow less obvious readings of a statute as a way of ensuring that something is done. While the assumptions underlying such decisions can certainly be criticized, it strains plausibility to do so by saying that Congress is always uniformly capable. Faced with situations that strike them as messy and unlikely to receive clear legal guidance, judges are more likely to adopt the attitude of “muddling through” (Lindblom 1959), allowing what strike them as pragmatic actions as long as they do not openly flout statutory text.
Section I of this chapter develops an abstract framework for considering the dynamics of “old law, new trick,” as well as providing (largely in the footnotes) many real-world examples to substantiate its claims. Sections II and III then contrast two of the highest profile episodes in which old statutes were asked to perform new tricks in recent years: the Clinton Administration Food and Drug Administration’s (FDA) attempt to use the Food, Drug, and Cosmetic Act to regulate tobacco, which was ultimately rejected by the Supreme Court in *FDA v. Brown & Williamson Tobacco Corp.*, and the attempt of various environmentalists and state attorneys general to force the Environmental Protection Agency (EPA) to combat global warming through regulation of greenhouse gases, which culminated in a victory over the reluctant agency in *Massachusetts v. EPA*. Close examination of these cases shows us the importance of institutional context, as well as giving a sense of the potential and perils of asking old laws to perform new tricks. Section IV offers concluding analysis, including further implications for theories of statutory interpretation.


Suppose policymakers have decided that some social ill is being inadequately addressed by current government policy—either because the problem is novel or because their priorities and judgments about which social problems are worth addressing differ from those of their predecessors. Policymakers can diagnose the existing shortcoming as an enforcement problem, an interpretive problem, or a legislative problem.

*Enforcement Problem:* Policymakers might decide existing laws and existing interpretations are entirely adequate to deal with the problem, but enforcement capacity or will is
lacking. To address such a problem, Congress could increase the appropriations dedicated to enforcement capacity at the agency, or it could pass a narrowly targeted amendment instructing the agency to place a greater priority on enforcing a specific provision of a law. The president or the administrator of the agency may also have the ability to reallocate resources to enforcement without a change in funding levels through various internal mechanisms. In the longer term, a president can seek to improve enforcement in a specific area by appointing personnel to the agency who share his priorities and will zealously pursue the goal in question (Lewis 2008).

Interpretive Problem: Alternatively, policymakers could decide that the substance of existing laws is adequate to address the problem in question, but the currently accepted interpretations of the laws have prevented them from being used to their full potential. The solution to such interpretive problems is straightforward: interpret some law differently so that it can legitimize the actions necessary to address the problem. This will be an especially appealing option when there is some broadly-worded statute that is germane to the issue at hand. Since the action of “interpretation”—or “reinterpretation”—can sound rather slippery, those pursuing this strategy are likely to brand their actions as merely “applying” the law as it is written. They are right to suggest that the line between a change in “enforcement priorities” and “interpretation” is a blurry one.

However the new use of a law is rhetorically branded, the obvious question is whether it is based on a reasonable, appropriate, and permissible reading of the existing statutory text. Opponents will argue that the problem is actually a statutory one, requiring legislative attention to address (if indeed it needs addressing). From their perspective, applying an existing statute to

---

4 See Davis (1969) for a classic treatment of how discretion is used in setting enforcement priorities, especially in prosecutorial judgment.
5 This is likely to be much easier if the agency’s external “constituencies” share the priority, or if the change is (or can be characterized as) incremental and building on the agency’s core competencies (Wilson 1989, 203-209, 221-232).
justify the contemplated action is an end-run around the proper legislative process, attempted precisely because there are not sufficient votes in Congress to support the new policy. If judges can be convinced that the new interpretation impermissibly strains the existing statutory text they will reject it and restore the interpretive status quo.

_Legislative Problem:_ Finally, there may simply be no law on the books that provides a means of addressing the problem—such is the downside of observing the rule of law.\(^6\) The clear solution is to utilize the constitutionally prescribed legislative process, complete with bicameralism and presentment. The benefits of this approach are clear, as the new legislative language will have the clear sanction of democratically elected officials, be as well adapted to the particular circumstances as is possible, and draw on the newest and best knowledge about how to address the problem sensibly. At the same time, the hurdles to passing new legislation amidst a crowded agenda are high, especially in the current age of polarization and procedural brinksmanship (Mann and Ornstein 2006, 2012).

How the problem is cast will have a direct bearing on which institutional actors are empowered to address it—and which actors may be rendered impotent without the cooperation of others. If we think of law as creating crystal clear requirements, the three categories might be thought of as mutually exclusive: enforcement problems represent executive branch enforcers “failing to do their job”; interpretive problems (which should be rare) represent interpreters, either executive or judicial, “getting the law wrong”; and legislative problems mean that changes in the law are entirely necessary. But if we think of law as frequently indeterminate, and “full

---

\(^6\) _Cf._ Letwin (1965, 101): “It seems clear, however, that a prosecutor who sought to stretch statutes beyond their ordinary meaning in order to prohibit the widest range of conduct…would be weakening the presumption of innocence and overly extending the power of government. Equipped with enough resources, such an officer might turn any modern society into a police state without invoking any authority beyond the already existing statutes.”
enforcement” as an idealized aspiration rather than a reality ever realized, then the three perspectives may be potential substitutes.

In some sense, any problem can be cast as a legislative problem. By reconfiguring or adding to the existing statutory edifice, especially through the clarification or further specification of the law, Congress has the power to alter the government’s orientation to nearly any social problem. If we imagine ourselves in the role of an omnipotent generic “policymaker,” as the opening question invited, pulling the congressional lever to change the law and thereby directly target the problem in question seems like the most straightforward approach to nearly any problem.

This option may not be as efficacious as it first seems, as substitutability may not be perfect. Change in the law will be not always be necessary or sufficient to secure the desired change in government action. As Chapter Five shows, if executive branch interpreters are committed to substantive ends in tension with the language of existing laws, their attempts to turn the law to their own purposes through “creative” interpretations will be difficult for courts to resist. Unless the change made to the law creates extremely clear requirements, the amendment may not accomplish its goals. On the other hand, if the problem really is an enforcement problem, Congress may need to support enforcement through increased appropriations, but it is unlikely to help the situation much by fixing an underlying statute that is not broken.

More to the point, in reality there is no generic omnipotent “policymaker” deciding which institutional mechanisms to use to address a problem. Rather, there are only government

---

7 There are certainly cases in which existing laws unambiguously offer options that have simply never been utilized (or have been underutilized) to that point in time. In that case, officials need merely to announce that they will be enforcing the law on the books more energetically than their predecessors. As an example, consider New York City Mayor Rudolph Giuliani announcing that various petty crimes previously overlooked—subway turnstile jumping, aggressive panhandling, prostitution—would be zealously enforced to the full extent of the law, which proved to be an effective way to target many criminals with outstanding warrants. See Kelling and Coles (1996) and Livingston (1997).
officials variously situated. Congress is a “they,” not an “it,” and a factious and overburdened “they” at that (Shepsle 1992). From the perspective of someone hoping to address a problem, asking Congress to amend their controlling statutes may be a first-best option, but the sitting legislature may well be unwilling or unable to act. Especially if there is divided government, a majority of legislators (or legislators in just one house of Congress) may be firmly opposed to the policy change in question, in which case passing a suitable amendment or new law will be impossible. Just as likely, though, is that some potential coalition of lawmakers exists to support legislative change, but that for any number of reasons it fails to coalesce.8

If Congress offers no immediate legislative solution, then, framing the problem as an interpretive one gives other actors an opportunity to change policy without statutory change. In other words, rather than being political entrepreneurs through effecting statutory change (Carpenter 2001; Sheingate 2003), various people may try to become legal entrepreneurs promoting new interpretations. Executive branch agency officials, who are the primary interpreters of most laws (Strauss 1990; Mashaw 2005), are in a position to effect policy change through a new interpretation. They were the sources of novel interpretations of the Sherman Act9, the Exchange Stabilization Fund created by the Gold Act of 193410, and the Food, Drug,
and Cosmetic Act. In doing so, they are likely to act as if there is no problem at all when one considers the text, legislative history, and underlying purpose of the relevant statute in the right light.

Novel interpretations may come from less centralized actors as well. Federal prosecutors may take the initiative in advancing a new interpretation as a way to increase the charges available to them in targeting wrongdoers. Such was the case in applying several criminal statutes to novel contexts, including the Refuse Act of 1899, the federal mail fraud statute, RICO and various Money Laundering statutes. Predictably, such prosecutorial efforts are applauded by the enforcement community and decried by the defense bar as causing over-criminalization where Congress never intended it. Private litigants are the other group capable of providing novel interpretations of old statutes, as they did in resuscitating the Reconstruction-era Ku Klux Klan Act for a variety of purposes, expanding the civil application of RICO, and

---

11 Agency officials provided the motive force for the new interpretation in the case of the FDCA and tobacco, as discussed in Section II.
12 33 U.S.C. § 407. Various prosecutors concerned with water pollution problems in the 1960s advanced an innovative reading of the Refuse Act, which had originally been passed for the purpose of keeping America’s waterways unobstructed for the purposes of navigability. Eventually, the Justice and Interior Departments under President Nixon also advanced this reading, which was ultimately superseded by the Clean Water Act of 1972 but had an important impact on that law’s architecture by changing the policy status quo that preceded it (Milazzo 2006, 164-170, 236-254). Rodgers (1971) applauded the novel trend in interpreting the Act, while Potter (1971) criticized the trend as inappropriately distorting the statute.
13 18 U.S.C. § 1341. Rakoff (1980) describes how prosecutors gradually turned the mail fraud statute into a general purpose anti-fraud statute, and Henning (1995) further describes the mail fraud statute’s various adaptation and suggesting limits to the scope of its jurisdiction.
16 Arguably, these represent cases of teaching an old statute the old tricks it was always meant to perform, but barred by courts from doing (Gressman 1952). In any case, it is certain that old statutes were dusted off and put to quite
pushing to use the Clean Air Act to combat global warming. In each case, the actors pushing the new trick hopes to advance their favored policy through having the status quo interpretation treated as the problem, rather than framing the problem as a statutory one susceptible only to a change in the law that may be harder for them to effect.

Not all instances of such thinking are likely to be controversial. There are probably innumerable instances of statutes being so adapted without ever bothering anyone. Just as, in the constitutional context, few originalists manage to get excited about the seeming unconstitutionality of the Air Force, many (and perhaps most) adaptations of static statutory text to a dynamic world are unlikely to ruffle many feathers.

However, when the reinterpretation in question creates distinct losers—often interests who will face more stringent regulation—it is likely to be challenged as an impermissible stretching of the statute. If an interest can show direct costs imposed by the new interpretation, it is likely to have standing to challenge the new interpretation in court. In Chapter Three, I argue that this is the main process through which legal constraints are made effectively binding.

extensive uses for purposes their drafters never anticipated. Achtenberg (1999) argues that § 1983 suits against state and local government officials who abused their positions are entirely consistent with legislative intent; Beermann (2002) chronicles the latter-day revival of § 1983 as the rights protected expanded, as well as the limited revival of §§ 1981, 1982, and 1985(3). Maine v. Thibotout (1980) allowed § 1983 suits to proceed to enforce Medicaid entitlements, which has had important effects (Mank 2003; Dunne 2007).

Lacovara and Aronow (1985) explain the expanded scope of civil RICO actions and criticize the underlying interpretation as contrary to the Act’s true purposes; Blakely and Perry (1990) take the opposite position, arguing for the propriety and manageability of these suits.

Private environmental litigants, joined by state attorneys general, were the ones demanding that the Clean Air Act be interpreted to include greenhouse gases. See the discussion in Section III.

Somin (2007) discusses this question and cites a number of scholars arguing that an independent air force ought to trouble strict originalists in spite of the apparent absence of originalists who are so troubled.

The case of the Exchange Stabilization Fund provides an instructive example of a situation in which nobody seems to have had standing to bring suit challenging the government’s novel interpretation. As a result, no case law exists about permissible interpretations of the underlying statute (Schwartz 1996; Henning 1999, 52-55). More generally, a sense of emergency is likely to lead courts to be more accepting of legally strained interpretations, as the alternative of waiting for the normal lawmaking process to play out seems less viable. One example, subsequently enacted into law, is the use of money laundering statutes, designed to retrospectively punish financial transactions used to support crime, for the purpose of prospectively freezing potential terrorist assets in the wake of the attacks of September 11, 2001. Gouvin (2003) recounts the interpretive changes advanced through an executive order by President Bush and expresses skepticism about the efficacy of using the statutes for this new purpose.
If there is someone to challenge the new interpretation, the question then becomes: what will lead judges to accept or reject the new interpretation?

Answering this question is not entirely different from trying to understand why judges make any particular ruling in cases of statutory interpretation. Judges will have to weigh the merits of arguments about what the text of the statute requires; what intent led Congress to enact the law, and what purpose it was meant to serve; and, if there is ambiguity, whether executive branch interpreters should receive deference. When evaluating new interpretations of old statutes, each of these familiar dimensions is likely to present itself in distinctive ways, and I consider each in turn.

A less familiar criterion, which I will argue plays a central role in judges’ decision-making in these types of cases, is a contextual empirical assessment of institutional competencies. Judges may share executive branch interpreters’ sense that the problem in question does indeed deserve new attention, but they may feel more inclined to hold out for the first-best option of fresh congressional action rather than supporting what may be an awkward reinterpretation of an existing statute. This inclination, though, will not be unlimited; where the contemporary Congress seems incapable of acting without some external stimulus, judges are likely to be more sympathetic to executive branch interpreters’ reinterpretation as a way to move policy forward.

A. Statutory Text

When a judge must decide whether a novel interpretation of a statute is permissible, the most important factor in her thinking will almost certainly be the plain language of the law—regardless of whether the judge thinks of herself as a committed textualist. The first question a
judge must ask when confronted with a novel interpretation of an old statute is: can the text of
the statute in question be reasonably read to support the action in question? Put conversely: is it
clear that the statute does not, on its face, rule out the action in question?

If they find that the text of the statute is clearly at odds with the interpretation proposed,
then judges will reject the interpretation as impermissible, except in unusual circumstances such
as when an absurd result would come from closely following the text (Manning 2003; Chapter
Two, Section 3.C). The very concept of having laws capable of exerting meaningful constraints
requires this, and the principle has also been famously codified in the seminal administrative law
case Chevron v. NRDC (1984), Step One of which says that if the language of the statute
unambiguously answers the relevant legal question, then the judge’s responsibility is always to
adhere to that meaning.21

In “old law, new trick” situations, the statute offered as support for the contemplated
action will almost certainly be at least potentially supportive of the action—indeed, the relevant
statutory text’s ability to support the novel interpretation and action is the whole reason it has
been chosen. On the other hand, if the old statute was designed for purposes other than the one
to which the new interpreters are seeking to put it, the statutory text is unlikely to resolve the
question at hand decisively in favor of the novel usage.

In many cases, however, the statute will have been chosen because by its plain language,
it does seem to include the proposed application and therefore support the novel interpretation,
and yet judges often balk at such seemingly straightforward applications of the statutory text.
They sometimes do so simply because they weigh textual factors less heavily than considerations
of statutory purpose or deference—each of which is discussed in turn, below. They may also do

---

21 The precise formulation is: “First, always, is the question whether Congress has directly spoken to the precise
question at issue. If the intent of Congress is clear, that is the end of the matter; for the court, as well as the agency,
must give effect to the unambiguously expressed intent of Congress” (842-843).
so, however, because they believe that the interpretation is only "seemingly" straightforward, and in fact represents a deliberate distortion of the statutory language when taken in its proper context. Judges reasoning in this mode may decry excessive "clause-bound" "literalism" as the enemy of sound textualism. If a judge feels that a particular word may only be stretched to include a novel application by ignoring the way in which that word is used within the context of the statute as a whole and in relation to connected sections, then she may try to reject a purportedly textualist interpretation as impermissible.

B. Statutory Intent and Purpose

If the statutory text is open-ended or ambiguous, such that it can plausibly be applied to the new circumstances, the judge’s next inquiry is likely to be whether the interpretation is also consistent with the underlying intent of the Congress that enacted the law and consonant with the purposes embodied in the act. In “old law, new trick” cases, it is unlikely that the new interpretation can be justified straightforwardly on the grounds that the enacting Congress specifically intended the act to be used in this new manner. If this can be persuasively established, there will be little cause for controversy. More likely, the argument will be that the enacting Congress deliberately framed the law in an open-ended manner so that its broad purpose could be realized in as-yet-unforeseen circumstances.

22 Legislative intent, and especially its representation in official legislative histories prepared by Congress, became a famously contentious subject in the 1980s and has remained a source of dispute. Textualist scholars and judges, prominently including future Supreme Court Justice Antonin Scalia and Judge Frank Easterbrook, drew on public choice scholarship questioning majority voting procedures’ ability to elicit a unique consensus to call into doubt the existence of any meaningful “intent” on the part of the legislature. For an account of these debates, see Farber and Frickey (1988), Eskridge (1988), and McCubbins, Noll, and Weingast (1994).

It should be noted that most judges and scholars who strongly oppose the use of legislative history do not despair in trying to understand the purposes of the enacting Congress. Rather, they merely believe that materials such as committee reports are unreliable guides to understanding these purposes, which they argue can be ascertained more reliably by turning to the text of the statute itself (Manning 2001).
Unless one takes the position that Congress is incapable of delegating the power to apply existing laws to novel circumstances—which only a few commentators are willing to do (e.g., Schoenbrod 1993)—then legislators must have the option to empower executive branch actors to recognize and address cognate situations as they develop. There is no question that legislators believe themselves to have this ability, as the usage of ambiguous and open-ended language is a frequently utilized tool in the creation of strong, forward-looking regulatory regimes. As a result, it may be possible to realize Congress’s purposes in enacting an open-ended regulatory law only by applying the law to new situations. The proponents of the new interpretation will certainly argue as much: that the new interpretation is faithful to the statute’s original visionary framers.23

One may fully support this style of reasoning and yet see its potential perils. Congress’s purpose in enacting any particular statute cannot have been to empower any part of the government to address every problem, or even to confer an unlimited creativity to meet every problem within a particular policy area. Almost no statutes take the form, “The agency may regulate problem X as it sees fit”; such a statutory abdication of responsibility to the executive branch would probably be found unconstitutional under the non-delegation doctrine, even given the decrepitude into which that strand of law has fallen. As a result, judges reasoning in a purposivist mode will focus on whether the open-endedness of the statutory terms in question justifies the novel application of the law.

They may decide that it does not for several reasons. The first is that the judge may believe that the enacting Congress would not have countenanced the proposed usage of the

---

23 For example, in arguing for the propriety of using the Refuse Act of 1899 to combat water pollution, Representative Henry Reuss (D – WI), Chairman of Subcommittee on Conservation and Natural Resources of the House Committee on Government Operations, declared that “The wise men in these seats in 1899 did it all…” (quoted in Potter 1971, 522n159).
24 Chapter Three’s reference to the Emergency Economic Stabilization Act provides a possible exception.
law—that the legislators at the time of enactment contemplated exactly such an application and
decided against it, even if such an application is clearly within the realm of reasonable
possibilities. In this scenario, the potential for applying an open-ended word to the new situation
is unintentional and indeed unfortunate given the compromise forged by the enacting Congress.
A judge can claim to understand the real content of this compromise either through legislative
history, other parts of the relevant statute, or contemporaneous enactments.

Supposing the enacting legislature did not specifically rule out the novel application of
the law, a judge may still reject the new interpretation because it “stretches the original purpose
past its breaking point.” In other words, following the new interpreters’ reasoning to its logical
end may have unacceptable implications in light of the purposes the statute was actually crafted
to meet.25 Although the particular application being pursued might plausibly seem to fit within
the original enactment’s purposes, accepting the new interpretation of the open-ended could have
the consequence of forcing acceptance of future actions predicated on identical logic that would
clearly be at odds with the statute’s purposes. Those opposed to the new interpretation will
invariably invoke a slippery slope argument, warning that allowing the broad interpretation is but
the first step toward allowing the law to be applied to nearly any situation. Such arguments
create an awkward situation for the new interpreters, especially if they are hoping to gain
significant new powers through their application of the old law. They must argue that they will
not seek to use the law as a pretext for an unlimited power grab while simultaneously claiming

---

25 For instance, in Bray v. Alexandria Women’s Health Clinic (1993), the Supreme Court refused an attempt to
extend § 1985(3) of the Ku Klux Klan Act to the protection of the suggested class of “women seeking abortions.”
Justice Scalia, writing for the majority, stated that if the court accepted such a claim “innumerable tort plaintiffs
would be able to assert causes of action under § 1985(3) by simply defining the aggrieved class as those seeking to
engage in the activity the defendant has interfered with. This definitional ploy would convert the statute into the
‘general federal tort law’ it was the very purpose of the animus requirement to avoid” (269).
that the assertion of power they are currently making is a reasonable one under the act. This
dynamic plays out in numerous “old law, new trick” scenarios.26

C. Ambiguity and Deference

If neither the statutory text nor an analysis of statutory purposes clearly resolves the
ambiguity, judges may be inclined to defer to the expert judgment of executive branch
interpreters, whose day-to-day experience in administering the statute arguably puts them in the
best position to apply the law to real world problems. Judicial deference to bureaucrats is a long-
standing trend in American administrative law, but, again, the rule is now most associated with
*Chevron v. NRDC*. Step Two of *Chevron* says that where a statute does not clearly resolve a
particular interpretive question, the executive branch’s interpretation should receive deference as
long as it is based on a “permissible construction of the statute” (843). In practice, Step Two
review of a statutory interpretation is often similar to arbitrary and capricious analysis under the
Administrative Procedures Act. To simplify, if an agency has taken pains to show why its action
is permitted under the statute and that there is also some rhyme or reason behind it, then judges
are likely to uphold the action.

In “old law, new trick” contexts, however, the question facing courts is somewhat more
complicated than normal matters of statutory interpretation. Rather than deciding on the
particulars of applying the law to a specific situation, the new interpreter is proposing to apply
the law to a novel class of cases, thereby expanding agency jurisdiction and substantive power.
This makes the normal rationale for *Chevron* Step Two deference problematic for three reasons.

---

26 For example, the argument over the proper application of RICO follows these lines precisely, with opponents of
the broad reading arguing that prosecutors inappropriately add RICO charges, creating the potential for larger
criminal penalties, whenever more than one person was involved in alleged criminal activity, while defenders of the
broad reading say that deterring white collar crime was among the real purposes of enacting the law (Goldsmith
2004).
First, whereas normally an agency can claim long experience of administering a statute as a source of relevant expertise, here the agency is attempting to break new ground. Although agency administrators can argue that their experience with the statute to that point has helped them to form a sound estimation of the consequences of applying the law in a new context, this claim is necessarily more speculative than when an agency is speaking from more direct experience. Second, for an agency to decide its own jurisdiction has the feel of being the judge in its own case—a situation which administrative law generally seeks to avoid. Opponents of the new interpretation are sure to make this point, alleging that deference in the case of a power grab is wholly inappropriate. Third, the idea behind deference to executive branch interpreters is not simply to empower bureaucrats instead of judges. Rather, *Chevron* deference is meant to facilitate Congress’s ability to delegate primary interpretive responsibility to agencies in specific contexts (Note 2005, 1689n7). Again, though, no act of delegation can permissibly be interpreted as wholly open-ended, so when the question is specifically about just how wide a particular delegation is, simply allowing an agency making reasonable arguments to decide the issue is problematic.

Textualist, purposivist, and deferential approaches to problems of interpretation are not best thought of as mutually exclusive alternatives. Judges normally take all of these factors into account as they attempt to determine whether an interpretation is permissible. The relative weight that a judge gives to each consideration may systematically differ from one judge to the next—this may be part of what is sometimes seen as judicial “ideology.” Just as important, however, judges are likely to give different weight to each of these factors in different contexts, including institutional contexts, which will be of special importance in “old law, new trick” scenarios.
D. Institutional Considerations

Whenever courts resolve interpretive dilemmas about statutory law they are compelled to decide, at least implicitly, about which institution will be the primary policymakers in the relevant area. Put in the terms offered above, in passing judgment on the permissibility of a novel statutory interpretation, judges will be determining whether the policymakers asserting the new interpretation were right to view their problem as merely interpretive, or whether in fact the obstacle to the new interpretation is actually a legislative one which can be remedied only by Congress. In neither case will the court be saying that the contemplated policy is off limits for all time, but it will be determining the limits of policy under the legal status quo.

Sometimes, this institutional consequence of judicial decisions may only be a by-product of judicial interpretations made wholly on the basis of other factors, like the ones discussed above. I argue, however, that especially in the context of “old law, new trick” interpretations, judges are likely to be keenly attuned to the institutional implications of their decisions, such that the policy-specific competencies of Congress and executive branch agencies will become important determinants of judicial opinions. Judges are fully aware that institutional competencies will largely determine the practical effects of their decision (e.g., Newman 1984). Classifying a problem as legislative and clearly making Congress responsible for driving policy change is likely to lead directly to new legislation in some contexts, while it may be met with legislative indifference or inefficacy in others. A judicial declaration that “this problem is legislative, making it Congress’s to fix, and we believe that Congress is more than capable of

27 This differs from the constitutional context, in which the Supreme Court’s substantive judgments about an issue are more like the “last word” on a particular interpretive question—though Whittington (2007) problematizes the idea of judicial supremacy in the constitutional context.

28 For a somewhat similar argument, cf. Barron and Kagan (2001, 212-213, 223), who argue, “Because Congress so rarely makes its intentions about deference clear, Chevron doctrine at most can rely on a fictionalized statement of legislative desire, which in the end must rest on the Court’s view of how best to allocate interpretive authority,” and assert that application of Chevron has always been responsive to institutional competencies.
addressing this problem through the legislative process\textsuperscript{29} relies upon an assessment of legislative capabilities that will be accurate in some situations and wildly Pollyannaish in others.\textsuperscript{30} If Congress has consistently proved itself to be too divided or distracted to address some problem with new focused legislation, judges may be more favorably disposed to permit a novel interpretation of an old statute, even if the fit of that old law is less than perfect.\textsuperscript{31}

In some cases, permitting such changes in interpretation will dramatically increase the chances that Congress will stir itself to act on a particular problem. Although the legal status quo prior to the new interpretation may have had the (at least passive) acceptance of a majority of legislators, the change brought about by the new interpretation may prod Congress to act in certain contexts.\textsuperscript{32} But judges may only want to give weight to such pragmatic considerations if

\textsuperscript{29} Probably the most common formulation of this idea is for judges to say that a problem is “for Congress, not the courts.” A Westlaw search of all federal cases in the last fifty years for this exact phrase yields 299 results, and of course there are many other ways to formulate the same idea.

\textsuperscript{30} Alternatively, if the text seems to support a broad interpretation, the court may tell litigants seeking a more restrictive interpretation that their problem is a legislative one. For one such declaration in the context of civil RICO, see Sedima, S.P.R.L. v. Imrex Co. (1985, 499), which held that although the statute was certainly being put to uses not intended by the enacting Congress, the situation “is inherent in the statute as written, and its correction must lie with Congress,” as courts would have no way of effectively fashioning a test to ensure that actions only be brought in cases of “real” organized crime and thus honored the original statute’s intent.

\textsuperscript{31} To be sure, the “vulgar formalist” (the term is described in the Introduction’s discussion as model I.B and comes from Leiter [2010]) who bases arguments entirely on a simple understanding of the relevant institutions’ formal competencies will reject this argument out of hand. The law is how it is, powers are allocated as they are, and that’s just the way the world is; we can’t wish ourselves some other sort of constitutional structure just because it seems to have led to bad consequences in one particular case.

That being said, most advocates of textualism try to make their arguments less dogmatically (since it is clear that if people reject the necessity of this perspective, simply insisting on its necessity won’t resolve anything). Few textualists—and especially few practitioners—are really willing to defend their position’s logical extreme, consequences be damned. Instead, they generally argue that textualism leads to the best system-wide consequences by preserving the legislature’s prerogative and disciplining judges.

The following hypothetical clarifies the dilemma facing the strictest textualist: there is some ambiguity about what the text permits, though the weight of the evidence suggests that it should not permit the action in question. However, if the new interpretation is not accepted, there is convincing evidence that disaster will ensue, with no possible timely legislative solution. Is it still the proper course of action to insist on interpreting the statute only in light of its framing, or is it appropriate in this case to also think about the consequences, and therefore accept the interpretation that seems less well-supported (though still at least somewhat plausible)?

If the strict textualist is willing to make a concession in this extreme case, the question simply becomes how clear and extreme indications of institutional deficiencies must be before they are willing to concede that textualism does not always trump other values.

\textsuperscript{32} This belief proved true in the case of using the Refuse Act to combat water pollution. Once judges accepted the new interpretation and thus changed the legal status quo, there was considerable pressure to create a powerful new
they believe legislative dysfunction is the likely alternative. If a textbook-civics response to the problem, characterized by congressional leadership, is likely to emerge without any judicial prodding, then judges are more likely to frown on strained adaptations.

Debate about questions of interpretation is almost never couched in exactly these institutionally sensitive terms. Instead, these institutional considerations enter when judges consider whether “Congress has spoken clearly to the issue.” Two parts of this inquiry can be potentially controversial. First, how clear is “clearly”? Second, how directly “to the issue” must Congress’s speaking have been? The clarity of a statute depends on the precision of its language. Roughly speaking, greater consensus and a clearer sense of congressional purpose will lead to more precise language. Language will be more precise when Congress was able to master the subject matter of the statute, agree on the means to be employed as well as the ends of the legislation, and win a strong consensus for a particular approach.

The functional precision of language is also likely to have something to do with the recency of the statutory enactment. In a changing world, new situations may arise that fit into the static classifications of a statute only awkwardly, and so, ceteris paribus, newer laws are clearer laws. When judges are tasked with determining whether an existing statute “clearly” resolves the interpretive question at hand, then, they are more likely to find that things are clear if a broad coalition of legislators has recently produced legislation on the subject. Additionally, there are epistemic reasons why newer text is likely to seem clearer to recent interpreters;

---

water quality law that would achieve many of the same substantive purposes without any of the policy awkwardness (Milazzo 2006).

33 Sunstein’s (1995) explanation of how ambiguity, even regarding basic purposes, is often a necessary ingredient for compromise provides some insight into why this is the case.

34 Eskridge and Frickey (1994, 57-8) offer a more practical reason why interpreters of a fairly clear recent law would want to hew close to the law’s clear meaning: “An interpretation in 1994 slighting the apparent meaning of a statute enacted in 1991 is likely to upset the coalition that produced the statute and, if the coalition is still powerful, subject the Court to the risk of a conflictual override.”
figuring out what a statute is supposed to mean is simply easier when the context of legislation is recent and thus easier to understand.\footnote{Easterbrook (1983, 534) notes: “Inferences [about ambiguous statutory meaning] almost always conflict, and the enacting Congress is unlikely to come back to life and ‘prove’ the court’s construction wrong. The older the statute, the more the inferences will be in conflict, and the greater the judges’ freedom.”}

Similarly, when deciding whether existing statutes speak directly enough to the issue at hand to be treated as dispositive, judges will have to consider institutional factors. If the legislative process is idealized as the univocal articulation of the general will, then the legislative body could be consulted and produce a clear pronouncement in every situation, but of course reality is quite different. Given a factious Congress with severely limited resources, agencies and courts must settle for statutes where the fit is “good enough.”\footnote{Recent congressional inaction is not a strictly necessary precondition for old law, new trick. The Refuse Act of 1899 was quite consciously invoked in favor of the more clearly applicable and less potent Water Quality Act of 1970. In sympathy with this maneuver, Rodgers (1971, 762) muses: “That the solons of the nineteenth century appear to have surpassed their modern successors in fashioning useful tools for combatting water pollution is a curious commentary on the accidents of legal history and on the vitality of the current drive to secure water quality.” As I argue below, however, the availability of a more recently produced statute is likely to make courts less sympathetic to the creative use of the old law.} Of course, in “old law, new trick” cases, this is precisely what is at issue: is the apparent relevance of the statute sufficient justification for the novel action? In deciding what fit will be considered “good enough” in these circumstances, judges will almost certainly be making an implicit comparison to the prospects of future legislation that will address the issue more squarely. Once again, if Congress has shown itself consistently capable of channeling political will toward a policy area in recent memory, then “good enough” may require clearing a fairly high bar. Judges will have to be convinced that applying the old law in question isn’t much of a stretch. On the other hand, if Congress seems capable only of dithering, then “good enough” may only require a plausible textual hook.

Once again, judges themselves do not generally reason using these institutional terms; rather, opposing camps thunder at each other, with one saying that, through the old law, Congress clearly covered the case at hand, and the other declaring that this is sheer nonsense. I
will argue that such heated exchanges, much in evidence in both *FDA v. Brown & Williamson* and *Massachusetts v. EPA*, mask deeper divides about the proper role of the courts in the context of a supine legislature. Strong textualists consistently make the case that Congress must be the driver of large scale policy change by arguing that the legal status quo “clearly” bars the interpretive innovation. Purposivists see broad statutes as giving the government broad responsibilities and are inclined to allow innovation, often by finding that the old statute “clearly” does apply to the new circumstances. As a close examination of the two cases will show, “clarity” in such cases is very much in the eye of the beholder.

Before finally moving on to examine the cases in detail, it is worth noting that if the proponents of the novel interpretation lose in court, their willingness to pursue a new interpretation may still represent a tactical victory for them by increasing the policy area’s salience and highlighting Congress’s recent inaction on the issue. Once a court decisively rejects their attempt to frame the issue in terms of an interpretive problem, proponents of policy change must frame it as a solely legislative problem, which clarifies Congress’s institutional responsibility. If the public was ultimately supportive of the goals of the novel interpretation, a loss in court may mobilize them and their representatives to seek the same goals through new legislative means. If the public was not ultimately supportive, perhaps they had little to lose.

Supposing the new interpretation is accepted as permissible, what are the benefits and costs for the proponents of the new interpretation? The benefits are clear enough: the new interpretation will lead to substantive policy changes, possibly far more quickly than they could have hoped to get any new legislation passed.\(^\text{37}\) There may be ongoing costs, though. Most speculatively, it is possible that by changing the policy status quo the impetus for fresh implementation of the new policy may be far from instantaneous. See Section III.

\(^{37}\) As the aftermath of *Massachusetts v. EPA* will show, however, if the drivers of the new interpretation are outsiders, the implementation of the new policy may be far from instantaneous. See Section III.
legislation could be diminished. This could be especially problematic in light of the policy awkwardness that is likely to be created by using an old law as the basis for novel actions for which it was not specifically designed. Such awkwardness may diminish the agency’s efficacy in achieving its policy goals as well as generating litigation—which can be used as a stalling tactic as well as a means of continuously re-opening the question of whether the new interpretation was really permissible, appropriate, and prudent. Finally, especially if the new interpretation survives the ruling of a divided court, there may be costs in terms of the agency’s legitimacy in the eyes of both the public and its “constituency” of regulated firms. Although successfully asserting its power might cheer sympathizers and represent a show of force, cultivating a reputation for unpredictability or needless provocation could ultimately damage an agency’s ability to effectively pursue its core mission (Carpenter 2010, 33-70).

To better flesh out our understanding of the dynamics of “old law, new trick,” in the remainder of this chapter I probe the details of two of the most prominent recent examples.

II. FDA and Tobacco

In this section, I explore the institutional and legal logic of the FDA’s assertion of jurisdiction over tobacco during the Clinton Administration, explaining the factors that led to the novel interpretation of the FDCA as well as its ultimate rejection by the courts. I then consider the subsequent policy history of tobacco regulation, including the passage of the Family Smoking Prevention and Tobacco Control Act of 2009, and reflect on the lessons for “old law, new trick.”
A. Background Context – Policy Responses to Tobacco and Historical Interpretations of the FDCA

Beginning in the 1960s, a broad-based cultural shift in the perception of smoking tobacco spurred policymakers to think about ways of mitigating the drug’s harms, including quite a few congressional enactments directly targeting tobacco usage. In 1965, Congress enacted the Federal Cigarette Labeling and Advertising Act (“FCLAA”), which required health warnings on cigarette packages, advertising, and billboards. The basic provisions of this law were reaffirmed by the Public Health Cigarette Smoking Act of 1969 and later by the Comprehensive Smoking Education Act of 1984, each of which made minor substantive amendments to the labeling requirements. Congress considered making the FDA responsible for these regulations but instead chose to empower the Federal Trade Commission (FTC) and Federal Communications Commission (FCC). The Alcohol and Drug Abuse Amendments of 1983 required the Secretary of Health and Human Services to investigate tobacco’s addictive nature and recommend appropriate action. In 1986, Congress passed and President Reagan signed a law requiring health warnings on smokeless tobacco packaging, to be administered by Health and Human Services. Finally, the Alcohol, Drug Abuse, and Mental Health Administration Reorganization Act of 1992 incentivized states to pass and enforce laws preventing minors from purchasing cigarettes by making mental health block grants conditional on these actions. Nevertheless, many believed that the regulatory regime jointly created by these programs was far too lenient on tobacco, leaving the public (and especially the young) with much too easy access to cigarettes and smokeless tobacco.

38 For a more detailed treatment of the legislative enactments mentioned here, including the full citations to each of them, see Brown & Williamson Tobacco Corp. v. FDA (4th Cir. 1998, 171-76).
39 Apart from issues related to regulating future tobacco usage, there was also the huge question of who should bear the cost of treating the many ailments caused by smoking. State attorneys general were most active on this front,
Running parallel to these many tobacco-specific congressional enactments, there was also a long history of interpreting the broadly-worded Food, Drug, and Cosmetics Act (FDCA). Given its modern form in 1938 (Pub. L. No. 75-717), the Act’s coverage of all drugs and medical devices certainly made it seem as though cigarettes might fall within the FDA’s jurisdiction. And over the years, federal courts had accepted FDA assertions of expanded jurisdiction over a wide range of products, even including a phonograph recording of a soothing voice “guaranteed” to put its listeners to sleep. The Agency had also managed to extend its reach over many novel medical developments without any statutory amendments to its charter, including “genetically modified foods, bioengineered drugs, nanotechnology, tissue engineering, and regenerative medicine, gene therapy, and pharmacogenics.”

At the same time, the FDCA makes no mention of tobacco, and the FDA repeatedly averred that it did not believe the Act included tobacco unless some particular manufacturer made health claims on behalf of their product. An official FDA Bureau of Enforcement pursuing claims that tobacco companies should be held responsible for the medical costs incurred dealing with tobacco-related illnesses. Negotiations meant to produce a “global settlement” involved Congress, though in the end the attempt to get a bill through failed, in part because of an inability to agree on the FDA’s role in tobacco regulation going forward (CQ 1998, 15-3-15-15). Eventually, the agreement between tobacco manufacturers and the states would be concluded without congressional sanction, and so of course included no definitive legislative resolution on the question of regulatory authority (Derthick 2002).

40 *U.S. v. 23, More or Less, Articles* (2nd Cir. 1951). Similar cases include *U.S. v. Halogenic Products Company* (D. Utah 1989), which sustained FDA jurisdiction to regulate a surgical instrument sterilizer as a medical device, in spite of its indirect connection to human health; *U.S. v. 25 Cases, More or Less, of an Article of Device* (7th Cir. 1991), which upheld the FDA’s assertion that a “sensor pad” claiming to improve women’s ability to conduct breast self-exams was a medical device under the FDCA; and *U.S. v. An Undetermined Number of Unlabeled Cases* (10th Cir. 1994), which upheld the FDA’s assertion that specimen-collection containers used by life-insurance companies were devices under meaning of Act, thereby creating jurisdiction. These cases all dealt with cases in which manufacturers made subjective claims about the effects of their products on human health, in contrast to the case of tobacco products. On the other hand, in *62 Cases, More or Less, Each Containing Six Jars of Jam v. U.S.* (1951, 600), the Supreme Court held that “imitation jam,” so labeled, was not misbranded under FDCA: “In our anxiety to effectuate the congressional purpose of protecting the public, we must take care not to extend the scope of the statute beyond the point where Congress indicated it would stop.”

41 See Noah (2008, 917-18) (sharply criticizing the agency’s tendency to view its organic statute as “a broad ‘constitution’ authorizing it to protect the public health by any necessary and proper means, rather than a limited and precise delegation of power from Congress).

42 For a more detailed description of these interpretations, including full documentation, see *Brown & Williamson Tobacco Corp. v. FDA* (4th Cir. 1998, 168-169).
Guideline in 1963 clearly stated that tobacco products did not fall under the act’s definition of a drug unless therapeutic claims were made on its behalf. In 1972, FDA Commissioner Charles Edwards testified in Congress that applying the FDCA to cigarettes would require their removal from the market, since it would be impossible to prove their safety, and again told legislators that the power to create future regulations was theirs, and not his agency’s.

Challenging the FDA’s disavowal of jurisdiction, an anti-smoking advocacy group petitioned the FDA to regulate tobacco in 1979. When the FDA demurred, the group sued, leading ultimately to a unanimous D.C. Circuit ruling in Action on Smoking and Health (ASH) v. Harris (1980). The court found that since the petitioners had no way of establishing, as required by § 201(g)(1)(C), that cigarettes were intended to have an effect on the body, the FDA’s interpretation of this section, which emphasized lack of claims by cigarette marketers (which was not rebutted by petitioners), was due deference. The judges concluded by insisting that “[i]f the statute requires expansion, that is the job of Congress” (243). This legal status quo was once again reinforced in 1989, when FDA Commissioner Frank Young reaffirmed for Congress the agency’s position that it was unable to regulate tobacco under its statute’s current form.

Throughout this period, Congress showed signs of understanding the FDA’s apparent lack of jurisdiction as it pursued its own tobacco-related agenda. Many bills were proposed to explicitly subject tobacco to regulation under the FDCA, but none ever reached the floor.43

Supporting the FDA’s longstanding position, though not addressing it directly, was the case of Federal Trade Commission v. Liggett and Myers Tobacco Company (D.D.C. 1952), in which an attempt to force the FTC to regulate advertising claims made on behalf of cigarettes was rejected, with the judge finding that cigarettes did not fall within the definition of drug under the FTC Act.

43 These include S 1468, 71st Cong. (1929); HR 11280, 84th Cong. (1956); S 1682, 88th Cong. (1963); HR 2248, 89th Cong. (1965); S 3317, 95th Cong. (1978); HR 279, 96th Cong. (1979); HR 3294, 100th Cong. (1987); S 769, 101st Cong. (1989); HR 5041, 101st Cong. (1990); S 2298, 102nd Cong. (1992); HR 2147, 103rd Cong. (1993).
B. The Novel Interpretation

At this point, President George H.W. Bush’s new FDA commissioner, David Kessler, entered the scene. Early in his tenure, Kessler decided to broach the question of whether FDA could use its statutory authority to pursue a more aggressive policy, but he found many agency veterans reluctant. These staffers worried that taking on tobacco would be costly, burn up the agency’s political capital, and make powerful political enemies. A younger cadre of idealists in the agency hoped for action, but Kessler’s initial sounding out of the agency led him to disavow jurisdiction (Kessler 2001, 35-6). For a little while, the issue was put on the agency’s back burner even as it received a steady stream of citizen petitions from tobacco activists (51). In the fall of 1992, an internal FDA study group on tobacco introduced Kessler to the idea of regulating nicotine as a drug under the FDCA (rather than figuring out a way to regulate tobacco more broadly), and it struck him as “a dramatic new way to approach an old problem” (63). Kessler was kept on by President Clinton, and soon his inclination to go after the tobacco companies was sharpened by the emergence of a whistleblower, code-named “Deep Cough,” who revealed that the tobacco companies were keenly interested in controlling nicotine levels even as they instructed employees never to discuss nicotine for fear that it could open them to FDA regulation (80-84). Kessler hoped to use such evidence to establish that tobacco manufacturers satisfied the intent requirement of the FDCA, even if cigarettes were not explicitly marketed as delivering a pharmacological effect. As the FDA’s investigation continued, Kessler sent Congress a letter indicating the agency’s changed position, hoping to “goad legislators into action” while FDA built its own case for unilateral action (88). Over the next two years the FDA conducted an extensive investigation, drawing increasingly strong congressional ire along the way (159, 163,
After convincing themselves that tobacco companies centrally conceived themselves as selling nicotine delivery devices, Kessler and other top FDA officials began to reformulate their interpretation of the FDCA, and specifically its requirement of manufacturer “intent” (270-2).

President Clinton only became actively involved in 1995, at which point dramatic action against tobacco presented one way for his administration to regain momentum after the Republican takeover of Congress in the 1994 elections (Kessler 2001, 322-4, 330-3). In August 1995, Clinton announced that he would be supporting “broad executive action” aimed to curtail youth smoking, and after the FDA finalized its rule in 1996 he trumpeted its importance in a Rose Garden ceremony (Kagan 2001, 2283).

The FDA’s proposed rule (1995a), which (including its jurisdictional appendix [1995b]) ran about three-hundred pages in the Federal Register, laid out the FDA’s youth-prevention-targeted policies: federal control of the minimum smoking age, prohibition of vending machines and free samples, and strong restrictions on advertising that could reach children or adolescents. The agency asserted jurisdiction over tobacco products not as drugs, but as medical devices—drawing an analogy between cigarettes and metered-dose inhalers. Their justification was laid out in a three part argument: 1) nicotine’s addictive and other pharmacological properties are effects on the “structure or any function of the body” within the meaning of the FDCA’s definition of a drug; 2) tobacco manufacturers intend their products to have these effects within the meaning of the Act; 3) regulation of cigarettes and smokeless tobacco products as devices is most appropriate at this time. The first two parts of this argument seemed to justify the FDA regulating nicotine as a drug, at which point the agency would be compelled to evaluate its safety.

---

44 Kessler would come to characterize this struggle (in the subtitle of his book) as “a Great American Battle with a Deadly Industry” and retells it as a sort of detective story.
and efficacy, but the third argued that the agency was entitled to regulate cigarettes as drugs, medical devices, or both, and that for prudential reasons the agency would choose to regulate them only as medical devices. The agency acknowledged that removing cigarettes from the market entirely would be inappropriate in light of the needs of 40 million addicted Americans.

The most radical interpretive shift was the move from the traditional, “subjective” understanding of intent, in which only manufacturer claims on behalf of a product were considered, to a new standard of “objective” intent, based on reasonable expectations about how a product would affect consumers (FDA 1995b, 44632-649, 44686-45204). In his account of events, Kessler therefore stresses that the FDA’s changed interpretation of the FDCA to apply it to tobacco did not rest entirely on a change of policy priorities, but rather was premised on the new information available to the FDA because of its investigations into the tobacco companies’ internal workings.45 Defenders of the FDA’s action asserted that this adjustment to new information is just how law should work in the context of the modern administrative state, especially for those parts of the executive branch charged with administering broad statutes. In an article defending the agency’s action, Cass Sunstein (1998, 1019) declared that “Without much fanfare, agencies have become America’s common law courts, and properly so,” given their superior ability to set out broad principles and then apply them to changing contexts.46

45 Gilhooley (2002) critiques Kessler’s strategy for wrongly fixating on the question of manufacturer intent and thereby losing track of congressional intent. As a result, she argues that Kessler needlessly led the FDA into a difficult legal position, where it might have been able to better stake claim over a narrower slice of tobacco policy related to youth prevention. Gilhooley nicely summarizes the courts’ dilemma in such cases: “The challenge of statutory interpretation is to determine when the agency’s resolution of a new issue is so far beyond the legislative aims that, even in light of Congress’s implicit delegation to agencies of the authority to adapt the laws they administer to new situations unless inconsistent with the specific provisions, the agency’s innovation should be found to be unauthorized” (1197-8).

46 Sunstein (1998, 1040) concedes that the original statute was not explicitly intended to cover tobacco, that the agency was not merely reacting to new information, and that it would be reasonable for opinions about applying the FDCA to tobacco to go either way. But he argued that principles of Chevron deference give good reason to defer to FDA’s judgment given its expertise.
Why did the FDA ultimately choose to pursue a strategy of “old law, new trick” in the case of tobacco? Kessler’s account emphasizes a genuine sense of conviction that the agency must do what was right, protect the public health from tobacco, and beat the tobacco companies, who conducted themselves so as to win Kessler’s undying enmity. He stood not in the place of a generic policymaker contemplating how best to engineer a public policy dealing with tobacco, but rather as an agency executive with certain tools at his disposal. Kessler seems to have recognized that being given new legislative tools might have been preferable, but his experience with Congress suggested to him that many members’ deep ties with cigarette manufacturers would make winning a legislative victory highly unlikely, especially after the Republican victory in the 1994 midterm elections. For his part, President Clinton was looking for opportunities to assert his continuing relevance after that setback, and Kessler’s enthusiasm for applying the FDCA to tobacco companies presented him with an appealing opportunity. Though a deal with the tobacco companies and their congressional allies might have been possible—especially on the issue of youth smoking prevention—neither Clinton nor Kessler was in the mood to patiently bargain and risk coming away with the policy status quo unaltered.

The opposition to the FDA’s new interpretation was swift and emphatic. Tobacco companies and their congressional surrogates decried the FDA’s legal imperialism and its disregard for the limits of its statutory authority. They emphasized the policy implication of applying the FDCA to tobacco, which is that cigarettes might not be long for the market given the statute’s requirements of medical safety and efficacy. Congressional opposition had geared up even before publication of the FDA’s proposed rule, with several bills being introduced by tobacco-state legislators to declare explicitly that the FDA’s actions were out of bounds.47

Opponents of the new interpretation also made their displeasure known during the notice and comment process of the FDA’s rulemaking. Apart from extensively challenging the science underlying the FDA’s findings about tobacco’s pharmacological effects and addictiveness, tobacco company complaints challenged the agency’s assertion of jurisdiction by questioning its interpretation of “intended to have an effect,” emphasizing that the FDA was dramatically reversing the agency’s earlier position, while simultaneously pointing out the difficulties of bringing tobacco under the FDCA’s regulatory requirements. The FDA universally rebutted these comments in its final rulemaking, making only relatively minor changes in the policy specifics of the programs proposed (1996a, 44417-26; 1996b, 44670-44685, 44995-45097, 45126-45204).

C. Judicial Reactions

The first court to hear a challenge to the FDA’s rule was the Middle District of North Carolina in *Coyne Beahm v. FDA* (1997). Judge William Osteen, a George H.W. Bush appointee, found that there was no statutory basis for FDA regulation of tobacco advertising, but he otherwise ruled entirely for the FDA, accepting the agency’s justification for interpreting the FDCA to include tobacco (Rienzo 1998). Citing *Chevron*, Judge Osteen declared that deference to the agency’s reasonable position was appropriate in light of the ambiguity in the statute, which gave no clear indication as to its applicability to tobacco.

The majority of a Fourth Circuit panel reversed in *Brown & Williamson Tobacco Corp v. FDA* (1998), calling out Judge Osteen’s opinion for having framed the issue as whether Congress

---

48 For an able rendering of all of these arguments, see Merrill (1998). Merrill was FDA Chief Counsel from 1975-1977 and was Lorillard Tobacco’s lawyer in the challenge against FDA’s assertion of authority. Merrill’s argument directly rebuts Sunstein (1998), criticizing Sunstein’s vision of agencies constrained only by explicit prohibitions as a radical rethinking of the structure of American government.
had clearly withheld jurisdiction over tobacco from the FDA. Instead, the inquiry under *Chevron*
should begin with whether Congress had ever evidenced any affirmative intent to delegate such
jurisdiction to the agency (161-62). The court found that the FDCA’s text did not provide
evidence for such a delegation unless one “examine[s] only the literal meaning of the statutory
definitions of drug and device,” and while the court admitted that tobacco seemingly falls within
these “literal” definitions, it insisted that the proper way of understanding the text was “in view
of the language and structure of the Act as a whole” (164). To show that the literal reading of
these definitions was misguided, the majority cited the many difficulties created by trying to
address tobacco under the FDCA’s requirements, concluding that the best evidence intrinsic to
the statute suggested that it was never intended to cover tobacco (167). The court also relied on
evidence extrinsic to the act itself to discern congressional intent, including legislative history,
the FDA’s historical stance against asserting jurisdiction, and, most pertinent to the discussion
here, a discussion of Congress’s actions pertaining to tobacco after the passage of the FDCA. In
a passage dealing with smokeless tobacco, the court declared that “the detailed scheme created
by Congress evidences its intent to retain authority over regulation of smokeless tobacco,” and
its logic was identical when discussing Congress’s various enactments regulating cigarettes
(175). The court concluded by insisting that the message of the case was that “neither federal
agencies nor the courts can substitute their policy judgments for those of Congress” (176).

---

49 For a criticism of the Fourth Circuit’s decision not to rely on *Chevron* and defer to the FDA’s interpretation, see
Sullivan (1999). For a defense, see Gellhorn and Verkuil (1999). They argue that *Chevron* deference should apply
only to interstitial gap-filling, and not to basic questions about the scope of an agency’s jurisdiction. They also offer
a rule of thumb: “The more significant the question and the greater the impact that expansion of the agency’s
jurisdiction is likely to have, the greater the likelihood that Congress did not intend implicitly to delegate that
determination to an agency,” at 1008. Other scholars (e.g. Sunstein 2006, 231-47; Moncrieff 2008) also believed
that *Brown & Williamson* effectively announced a “major questions” exception to *Chevron*. However, I argue
below that we learn from the Supreme Court’s decision in *Massachusetts v. EPA* that importance is not necessarily
the critical element compared to legislative assertiveness.
A five-member majority of the Supreme Court closely followed the Fourth Circuit majority’s logic in *FDA v. Brown & Williamson Tobacco* (2000), again finding the FDA’s novel interpretation to be impermissible. Justice O’Connor, writing for the majority, bent over backwards to emphasize the seriousness of smoking as a social problem worth addressing, in her very first sentence labeling smoking as “one of the most troubling public health problems facing our Nation today” (125). Nevertheless, the majority once again found that the intrinsic evidence of the statute’s true meaning “clearly” precluded covering tobacco, if one avoids the pitfall of “examining a particular statutory provision in isolation” (132). O’Connor relied more heavily on the difficulties of reconciling tobacco’s effects with the FDCA’s requirements of “safety,” saying that “if [tobacco products] cannot be used safely for any therapeutic purpose, and yet they cannot be banned, they simply do not fit” (142-3). And, like the Fourth Circuit, she heavily emphasized the evidentiary value of repeated congressional enactments specifically addressing tobacco. The creation of “a distinct regulatory scheme” for tobacco, which at all times was informed by the FDA’s disavowal of jurisdiction, made it impossible to believe that the question was still an open one (157).

---

50 O’Connor was joined by Rehnquist, Scalia, Kennedy, and Thomas; Justice Breyer, in dissent, was joined by Stevens, Souter, and Ginsburg.

51 O’Connor does not follow the Fourth Circuit in decrying “literalism,” but the tenor of her opinion’s discussion about interpreting the FDCA’s definitions is much the same. Manning (2000, 256) argues that the spectacle of these avowedly textualist Supreme Court justices wriggling out of a straightforward reading the language at issue shows that the Court was willing to privilege non-delegation concerns, which suggest that such a large transfer of regulatory authority without painfully clear evidence of explicit legislative intent is problematic. Manning believes that the Court’s adherence to the doctrine is fundamentally misguided, though, and ultimately “creates the perverse result of attempting to safeguard the legislative process by explicitly disregarding the results of that process” by substituting judicial lawmaking for legislative. He argues that judges should respect Congress’s ability to leave future resolution of important questions to agencies through broad statutes.

52 Elhauge (2008, 102-104) asserts that the majority’s central holding, viewed correctly, was the rejection of the FDA’s position that its new regulations reflected a legitimate instantiation of current political preferences—or “currently enactable preferences,” in Elhauge’s terminology. For him, the recent congressional activity regarding tobacco is most notable for its ultimate failure to produce any decisive statement on FDA jurisdiction, which suggested that a move toward dramatically expanded FDA jurisdiction was not currently enactable. This seems to get things precisely backward, though. The FDA’s problem was not being out of touch with the public mood—indeed, later developments (discussed below) give good reason to think it was doing a fair job reading the political winds. Rather, it is that there are good reasons to reject such a novel application of the FDA’s existing statute when
Justice Breyer penned a sharp dissent, as well as taking the unusual step of reading portions of his opinion from the bench (Lash 2000). As a starting point, he emphasized the literal applicability of the FDCA’s definitions as well as the FDCA’s broad scope and purpose of protecting public health, both of which suggested the propriety of the FDA’s interpretation. Given these facts, Breyer suggests it is hard to take seriously the majority’s finding that the statute 

*clearly* 

excludes tobacco. He then rebuts the majority’s finding that tobacco could only awkwardly be fit into the statute’s framework, arguing that if the definitions fit and the intent to affect the body has been clearly established, then the FDA has jurisdiction, and the question of how to fashion the most effective remedy is left to the agency. Therefore, although the statute could be used to support a ban of tobacco products, the agency’s decision to proceed with less draconian measures—in part out of a concern for the black markets that would arise to supply smokers’ demands—must not be ruled unlawful simply for being prudent (174-181).^53^ Meanwhile, Breyer argued that the tobacco-specific legislation enacted after the FDCA unvaryingly failed to include any explicit declarations about FDA jurisdiction, and so whatever presumptions about that jurisdictions legislators may have had, their actions should be regarded as having left FDA jurisdiction untouched (181-86).^54^ 

---

^53^ Breyer’s thinking here is remarkable, suggesting as it does that executive branch agencies are permitted to act on their own expectations of perverse consequences, Congress’s direct instructions to the contrary notwithstanding. Following this logic, any time executive officials believed legal requirements would produce negative net social consequences, they would be justified in setting aside the law in the name of the public good. This more closely resembles the hoary doctrine of royal prerogative (e.g. Locke 1980 [1690], ch. XIV) than it does any accepted doctrine of modern American administrative law. Breyer justifies this position by briefly asserting that the issue’s high profile would assure that the public would associate the policy with the president, thereby mitigating any problems of democratic accountability; 529 U.S. at 190-191. Once again, a few sources support this plebiscitarian view (e.g., Mashaw 1985), but most scholars would see cutting the people’s representatives in Congress out of the loop as raising deep constitutional issues. Wallach (2010) provides a deeper exploration of these issues in another context.

^54^ Manning (2000, 264) agrees with Breyer that the Court’s argument is quite weak, since “enacting a statute based on an assumption about law does not amount to enacting that assumption.” If, as I suggest, the recent enactments are relevant as evidence about institutional competencies rather than as indicia of the FDCA’s purpose and meaning,
D. Responding to the Supreme Court’s Ruling

In the immediate aftermath of the Supreme Court’s ruling, the FDA made provisions to quietly wind down and end the programs that had been set in motion under its tobacco rule (FDA 2000). Some groups petitioned the agency to attempt to make a new attempt to regulate the marketing of those cigarettes whose manufactures claimed containing fewer toxins than competitor brands’, since such claims had served as the basis for FDA jurisdiction in the distant past (Fairclough 2001). The agency took no immediate action on these petitions, though.

While it ended the executive-branch-initiated program to regulate tobacco under an old law, the Supreme Court’s ruling also spurred a great deal of political agitation for a new law tailored to address tobacco. Signs of consensus seemed promising, especially on the narrow issue of youth smoking prevention. Responding to the Supreme Court’s decision, President Clinton called on Republicans to join with him in curbing tobacco sales and marketing and endorsed a bill sponsored by Senators Bill Frist (R – TN) and John McCain (R – AZ) (S. 2566, 106th Congress); presidential candidate and Vice President Al Gore called on his opponent, Texas Governor George W. Bush, to join him in supporting the effort to empower the FDA to regulate tobacco (Otto 2000). Bush responded by also calling for increased regulation of tobacco this point becomes less relevant. Manning goes on to argue that treating the recent enactments as controlling, such that they should not be implicitly repealed by a novel interpretation of an older and more general statute, is a more sensible argument for the majority. He views such a move as a proper application of the principle of “the specific controls the general,” which he believes better promotes values of accountability and democratic process than does the non-delegation doctrine (276). To put this canon in the language of this essay, all one has to do is read “the specific” as recent targeted legislative actions and “the general” as older, broadly-worded statutes. Their claims were based in part upon the fact that the FDA had successfully, and uncontroversially, asserted its jurisdiction in two cases in which cigarette manufacturers marketed their products as weight-loss aids in United States v. 46 Cartons, More or Less, Containing Fairfax Cigarettes (D.N.J. 1953) and United States v. 354 Bulk Cartons, More or Less, Trim Reducing-Aid Cigarettes (D.N.J. 1959).
Surprisingly to many, the tobacco companies themselves were also softening their position by admitting the need for some regulation. The giant Philip Morris corporation, in particular, had already begun to welcome anti-youth smoking measures even before the ruling. For all this talk, the bill never made it out of the Health, Education, Labor and Pensions Committee in the Senate. Similar efforts garnered bipartisan support in 2001, but the bills (H.R. 1043, H.R. 1097, H.R. 2180) once again died in committee in the 107th Congress. George W. Bush’s presidency saw serious action taken on a bill in 2003, with Philip Morris again supporting the effort. Senators sought to fashion a compromise fusing a bill to give the FDA jurisdiction with a bill to end the New Deal-era tobacco subsidy program, but smaller tobacco companies’ opposition to an expanded FDA role—largely because they felt the new, more-regulated regime would help to lock in Philip Morris’ market dominance—helped to doom the attempt that year (CQ 2003, 11-18-11-19). A similar drama played out in 2008, when a bill to give the FDA jurisdiction over tobacco passed in the House, but Senator Ted Kennedy’s (D – MA) bill in the Senate never received a floor vote due to various camps in opposition (CQ 2008, 3-14-3-16).

Finally, the Family Smoking Prevention and Tobacco Control Act (FSPTCA) was enacted in June 2009 (Pub. L. No. 111-31). Though some Republicans championed an alternative that would have created a new agency for regulating tobacco, in the end advocates of FDA jurisdiction won the day (CQ 2009, 17-4-17-6). The law created a dedicated Center for

---

56 Many leading Republicans voiced their doubts about entrusting such power to the FDA. E.g. House Majority Leader Tom DeLay (R –TX), who declared, “I oppose legislation that would expand the bureaucratic reach of the very agencies so intent on circumventing the role of Congress on important matters of public policy” (reported in Lash 2000).

57 Philip Morris began actively positioning itself as favoring policies to prevent youth prevention of smoking even before the Supreme Court ruled on Brown & Williamson—though, even as it admitted the addictive nature of nicotine, it was always insistent that it was improper to classify tobacco as a “drug” under the FDCA (“Verbatim; Big Tobacco's Changing Tune,” 2000; Flanigan 2001).

58 Once again, opponents included the smaller tobacco companies, but at least a few Republican Senators claimed to ground their opposition on a desire for a stricter bill.
Tobacco Products within the FDA to receive its own budget and administer specially tailored rules with different requirements than the “safety and efficacy” required by the FDCA (§ 901(e)) (Curfman, Morrisey, and Drazen 2009). In what must have been a satisfying moment for Kessler and Clinton, the FSPTCA specifically required the FDA to reinstate the substance of the 1996 tobacco rule’s youth smoking sections (§ 102(a)(2)) (Termini 2011). Crucially, though, the FSPTCA also contains an explicit limitation on the FDA’s powers: the agency may not ban cigarettes or require nicotine levels to be reduced to zero (§ 907(d)(3)).

E. Analysis

What does the FDA’s attempt to assert jurisdiction over tobacco under the FDCA teach us about the dynamics of “old law, new trick”?  

First, in the choice to adapt an old law in the first place rather than focusing on passing new legislation, we see the importance of considering policymakers not as an undifferentiated, unitary body, but rather as an aggregate of individuals situated within particular governmental institutions. Kessler’s FDA acted on its own initiative, and then won the support of the president, quite independently of Congress. It chose to adapt the FDCA to its present purposes because doing so provided a promising opportunity to empower itself—but there is no reason to view this decision as cynically motivated. Rather, the people within the FDA had ideas about their role and their agency’s mission that pushed them to embrace responsibility for regulating tobacco. 59 Although in the presence of opposition, they recognized that their action would have

---

59 Carpenter (2010) provides much more on the sense of mission within the FDA and its relationship to the agency’s external reputation. Carpenter argues that Kessler attempted to use the FDA’s reputation as a “gatekeeper” to justify its jurisdiction, even if he “knew he was appropriating a legal, political, and conceptual architecture established for other purposes” (745). Carpenter (briefly) frames the issue in terms of reputation; here, and in general, this useful outlook does not always give much consideration to the ways in which the agency remains a creature of law and particular statutory authorities and limitations.
the appearance of teaching the FDCA a new trick, they sought to portray the move as growing organically out of their agency’s mission and legal obligations. Clearly the new interpretation would lead to a large expansion of agency power, but their conviction was that the law was on their side regardless of what past interpreters thought about tobacco.

Why did the courts ultimately decide that this new trick was impermissible? The opinions of the 4th Circuit and Supreme Court majorities lay out three main reasons for the judges’ rejection of the FDA’s position. First, there is the policy awkwardness created by attempting to fit tobacco into the FDCA’s provisions. The judges were certainly right that incongruities and difficulties would exist—but if the statute decisively required regulating tobacco, it seems hard to believe that the judges would take it upon themselves to fashion such a monumental exemption. Many statutes apply awkwardly to new situations presented by a changing world (apart from any “old law, new trick” situations), and courts frequently express the idea that Congress must be the one to provide amendments to diminish the awkwardness.

Second, and related, is the assertion that the FDCA clearly precludes any jurisdiction over tobacco. From this angle, the awkwardness is evidence that the statute’s true nature is incompatible with regulating tobacco, an argument pursued by examining the structure and history of the FDCA. Here, Justice Breyer and the dissent have the better of the argument: the Act’s broad language seems to encompass tobacco and it was Congress’s choice to make the statute a broad and flexible one. At the very least, this makes it hard to accept that the statute clearly precludes regulating tobacco. Of course, the majority is probably correct that it is “clear” that Congress never specifically anticipated FDCA applying to tobacco as it contemplated enacting the FDCA, but such a fact is quite beside the point in trying to understand a broad statute. Requiring that sort of foresight defeats Congress’s ability to create flexible authority
capable of responding to new information. And information really had changed—Kessler’s FDA put together very impressive evidence on manufacturer intent that had never before been publicly known.

The third and strongest argument offered by the majority rests on their citation of the many tobacco-specific congressional enactments over the years, all enacted in the shadow of the FDA’s disclaiming any authority over tobacco under the FDCA. Although the majority sometimes talks about these as if they speak directly to the question of whether the FDCA can be applied to tobacco, this seems to be something of a stretch—the working assumptions of legislators do not in any way become a part of the law unless they are given an anchor in the statutory text, as they were not in this case (Manning 2000). A far more convincing way to understand this evidence is to imagine executive creativity as a practical substitute for legislative action, at least in situations where the legislature has conferred broad authority in a relevant area. As discussed in Part I, this way of thinking posits “normal” lawmaking, complete with bicameral approval and presentment to the president, as the first-best, most preferred means of policymaking, but recognizes that applying this model to every issue that arises may be practically impossible. Where legislative action is not forthcoming, either because issues are not salient enough to garner congressional attention or because Congress proves itself incapable of action on a particular issue, courts are then likely to be indulgent of bold legal interpretations to support novel executive branch actions. In the case of tobacco, the court viewed Congress as capable of leading the development of policy itself. The citations of Congress’s many tobacco-specific laws provided strong support for this position, which the subsequent passage of the FSPTCA proved to be justified. Given such robust congressional activity on this policy area, no executive-initiated “old law, new trick” substitute was needed, or warranted.
From this perspective, the idea that subsequent events showed that the court erred in *FDA v. Brown & Williamson* seems quite nonsensical. Although the final law did reinstate the 1996 law that the court struck down, in many ways the 2009 Act was not a close substitute for an FDA victory in the earlier case. Had the FDA prevailed, it would be difficult to find any legal (and not merely prudential) reason why the agency would be constrained from laying waste to the legal trade in tobacco at any time. Indeed, such a remedy would have always been the most natural way of applying the FDCA to tobacco, which is unlikely ever to be “safe” in the sense required by that law. Such an outcome would undoubtedly have embroiled the agency in years of bitterly contested litigation as well as making it the target of political backlash. The FSPTCA, on the other hand, enacted a more workable compromise, giving the agency tobacco-specific powers and responsibilities and clarifying that its actions are not to effectively promulgate any ban or requirement that nicotine levels be reduced to zero. The law, therefore, represents a far more sustainable and coherent foundation for ongoing policymaking efforts.

Given the ultimate outcome, it is certainly possible to think that the FDA’s attempt to teach an old law new tricks ultimately paid dividends in terms of policy change, even if the new trick was itself rejected. Although imagining the counterfactual world in which the FDA never took its case to court is quite difficult, the FDA’s novel interpretation of the FDCA may have been the crucial ingredient in jump-starting meaningful policy change. In his book, former FDA Commissioner Kessler (2001, 383-85) took this stance. He lamented that the Court’s majority had been “unable to recognize how much had changed” to justify the change in the FDA’s interpretative position, and concluded that the majority had simply followed its “ideological…attitudes toward government regulation” in voting against his agency.

---

60 Admittedly, even under the FSPTCA there are some outstanding free speech challenges to the FSPTCA concerning the constitutionality of its warning label requirements and advertising restrictions (Selyukh and Pelofsky 2011).
Nevertheless, he was heartened by what he perceived as a changed tone of debate regarding tobacco, signified by Philip Morris’ public reversal and the large punitive damages that juries were beginning to assess against cigarette manufacturers. Knowing how policy has developed since, his assessment of the FDA’s impact seems quite defensible.

III. The Clean Air Act and Greenhouse Gases

In this section, I examine the push by various environmentalists and states to compel the EPA to classify greenhouse gases (GHGs) as pollution agents so that they would be covered by the Clean Air Act (CAA). These parties litigated their claim against a resistant Environmental Protection Agency, eventually triumphing with a narrow Supreme Court majority ordering the agency to cover GHGs under the CAA in *Massachusetts v. EPA* (2007). I consider the ongoing response to the ruling and analyze why the Court ultimately ruled in favor of “old law, new trick” in this context, in many ways similar to *Brown & Williamson*. Once again, I argue that institutional considerations provide the best grounds for understanding the Court’s decision.

A. Background Context

The public gradually became aware of, and concerned about, global warming in the 1980s and 1990s. As the science establishing global warming became better established, recognition of global warming and support for government action to address it both grew steadily through the early 2000s (Nisbet and Myers 2007), though during the past decade public opinion has become increasingly polarized (McCright and Dunlap 2011, 155, 176-8).

61 The divergence between educated Democrats and Republicans is especially notable; nevertheless, levels of concern remain higher than they were in earlier decades.
Congress had responded to global warming concerns by sporadically funding research on the subject.\textsuperscript{62} The National Climate Program Act of 1978 required the creation of a research program which eventually issued a report stating that current trends seemed to be leading to significant warming (Climate Research Board 1979). The next action came with the Global Climate Protection Act of 1987, which requested that the EPA formulate a response plan to global warming. President George H.W. Bush signed, and the Senate unanimously ratified, the 1992 United Nations Framework Convention on Climate Change (UNFCCC), a nonbinding agreement to work toward preventing and mitigating the damages of global warming. When the UNFCCC convened again in 1997 in Kyoto, Japan, the resolution they eventually approved would have bound the United States and other developed economies to make significant reductions in their GHG emissions, while simultaneously creating much more modest requirements for developing nations. Disfavoring these terms, the U.S. Senate passed a resolution, 95-0, expressing its disapproval of the emerging Kyoto treaty on July 25, 1997 ("The Senate Vote on a Climate Treaty," 1997). The Clinton Administration, especially Vice President Al Gore, remained ardent supporters of the Kyoto Protocol that was eventually approved by the Convention, signing it in late 1998 (Cushman 1998), but they recognized that they had failed to secure a treaty that the U.S. Senate would approve, and so never submitted the treaty to a vote (Rabe 2004, 15). Following Kyoto, Congress showed considerable interest in global warming issues, with members introducing hundreds of bills in the late 1990s and early 2000s, though none led to legislation (Moncrieff 2008, 636). In total, then, legislative action addressing global warming remained quite limited, especially in comparison to the statutes regulating tobacco.

\textsuperscript{62} For a full discussion of these congressional actions, include full documentation, see the majority opinion in \textit{Massachusetts v. EPA} (2007, 507-9).
During this period, the EPA was also mostly quiet on the global warming front. Still, in two instances in which the agency was asked to express its opinion about whether it was capable of addressing global warming given its current set of policy tools, it indicated that its interpretation of the CAA gave it such a power. At a hearing in March 1998, EPA Administrator Carol Browner had testified to Congress that the agency already possessed jurisdiction to regulate carbon dioxide (CO$_2$). In response, Congressman Tom DeLay (R – TX) asked the agency to produce a legal justification for this position. This led to the so-called “Cannon Memorandum” (EPA 1998), in which EPA General Counsel Jonathan Cannon briefly outlined for Administrator Browner the basis for asserting EPA jurisdiction over GHGs. In just a few paragraphs, Cannon explained that the definition of “air pollutant” found in § 302(g) of the CAA$^{63}$ seems to clearly encompass CO$_2$, with its natural occurrence having no bearing on whether it can be a pollutant at some level, and that the EPA therefore has jurisdiction to regulate power plants’ CO$_2$ emissions if it makes a finding that CO$_2$ endangers public health or welfare (EPA 1998, 3). He then noted that the EPA has made no such finding, nor has any specific plans to do so, but that the requirements for regulation “could be met” if the Administrator determined that harm could be “reasonably anticipated” (4). Congress followed up with a hearing devoted specifically to this question, at which new EPA General Counsel, Gary Guzy, reaffirmed the agency’s belief that it had the power to regulate CO$_2$ (EPA 1999). The Clinton Administration never did take the decisive step of asserting jurisdiction before it ended, though.

---

$^{63}$ The section states in full: “The term ‘air pollutant’ means any air pollution agent or combination of such agents, including any physical, chemical, biological, radioactive (including source material, special nuclear material, and byproduct material) substance or matter which is emitted into or otherwise enters the ambient air. Such term includes any precursors to the formation of any air pollutant, to the extent the Administrator has identified such precursor or precursors for the particular purpose for which the term ‘air pollutant’ is used.”
B. The Novel Interpretation

Instead, unsatisfied with congressional actions, and hoping to convert the Cannon memorandum’s nonbinding legal judgment into action, various environmental groups petitioned the EPA to regulate GHGs under the CAA in 1999. The petitioners, who would come to include twelve states, three cities, and one territory in addition to the environmental groups, asked the EPA to regulate mobile sources’ (mostly automobiles) greenhouse gas emissions under §202(a)(1) of the CAA, which requires that the EPA shall “by regulation prescribe…standards applicable to the emission of any air pollutant from any class or classes of new motor vehicles or new motor vehicle engines, which in his judgment cause, or contribute to, air pollution which may reasonably be anticipated to endanger public health or welfare.” The key term “air pollutant” is defined by CAA § 302(g) to include “any air pollution agent or combination of such agents, including any physical, chemical, biological, radioactive … substance or matter which is emitted into or otherwise enters the ambient air.” The petitioners argued that this section should certainly be interpreted to include GHGs, since, in their view, these emissions' contribution to global warming was clear enough to make designating them as “air pollution agents” entirely natural.64

Once again, we see that the impetus for pursuing a novel interpretation came from participants in the policymaking process who did not possess the leverage to push through legislative change. Although environmental groups are certainly not without staunch congressional allies, on the issue of global warming in the late 1990s, they (probably correctly) understood that there was not a sufficiently large political consensus to address what might most naturally be treated as a legislative problem. They also managed to enlist a number of state

64 More specifically, petitioners argued that various EPA statements made on its official website and in other agency documents provided sufficient evidence of the connection between GHGs, global warming, and potential harms to the public health and welfare to justify treating GHGs as pollutants (EPA 2003b, 52923).
attorneys general to their cause—another group of actors who has sought to shift policy through litigation at the national level (Nolette 2011, 296-300; Provost 2010).65 Both of these groups saw an opportunity to present the lack of GHG regulation as an interpretive problem rather than a legislative one. Similar to the case of the FDCA and tobacco, they argued that the Clean Air Act is a classic example of a broadly worded statute designed to include new substances or threats as they become scientifically established. Like the Clinton FDA, they therefore took the position that the novel interpretation they were insisting upon was not really a “new trick” at all—rather, it was simply a logical reading of the obligations the CAA creates for the EPA.

In this case, however, advocates of a new reading did not have the executive branch agency as an ally. The Clinton administration EPA’s cautious expression of support for asserting EPA jurisdiction gave way to a Bush administration EPA that struggled with the question of legal authority, but ultimately chose to deny the petitions. The agency charged with administering the regulatory statute thus became the most important opponent of this particular attempt to teach an old statute a new trick.

The EPA, first in an internal memorandum (2003a) and then in expanded and formalized form in a Federal Register notice (2003b), argued that the CAA did not confer the authority to deal with global climate change—especially under the precedent of Brown & Williamson v. FDA. Citing the plain text of the CAA, the EPA pointed out that the Act contained no explicit

65 As Nolette (2011) describes, the State AGs originally took a different approach in their suit against the EPA, which was originally filed as Massachusetts v. Whitman in the District of Connecticut in 2003. This original complaint attempted to force the EPA to list CO₂ as a criteria pollutant under § 108, which would automatically trigger a number of statutory requirements to regulate emissions from stationary sources, but was withdrawn after the EPA officially denied the environmental groups’ petition and the State AGs decided that they would be better off joining them in suing over the § 202 denial. Nolette also describes a parallel “old law, new trick” strategy pursued by the State AGs as the litigation against the EPA proceeded. In this case, the law was common law—specifically, the doctrine of public nuisance, which the state AGs proposed to turn against various electric utilities for their contribution to global warming. This approach was ultimately rejected by a unanimous Supreme Court in American Electric Power v. Connecticut (2011). Justice Ginsburg’s opinion more or less said that there was enough action on other fronts to make entertaining such an ambitious common law challenge superfluous, fitting in nicely with the overall argument being advanced here (Nolette 2011, 301-6, 322-330).
language encouraging regulation targeted at global climate change, and in fact such a provision was considered as a part of the 1990 CAA Amendments and subsequently left out of the act (2003b, 52925-26). Given that the only explicit mentions of GHGs or CO₂ in the CAA pertain exclusively to research⁶⁶ and an assertion of agency authority would have profound economic consequences, the Court’s majority decision in *Brown & Williamson* suggested that the agency should not look to drastically expand the scope of its regulatory authority without more explicit congressional sanction (52928). This was especially so given various mismatches between the CAA’s locally-oriented structure and the global nature of climate change (52927).

Much like opponents of applying the FDCA to tobacco, the EPA also cited other relevant government efforts to effectively regulate GHGs already underway under other authorities. Noting Congress’s various efforts to spur GHG research, the United States’ participation in the first UNFCCC, as well as assorted failed attempts to legislate more ambitious policies, the EPA argued that “this backdrop of consistent congressional action to learn more about the global climate change issue before specifically authorizing regulation to address it” made it clear that EPA was not permitted to simply cover GHG emissions under the CAA by a change in interpretation (52923-26). Relatedly, the EPA pointed to the likely redundancy of regulating automobile CO₂ emissions given already existing fuel economy standards enforced by the Department of Transportation (52929). The EPA argued (and proponents of the novel interpretation denied) that the totality of these actions demonstrated Congress’s ability to sensibly steer the development of climate change policymaking, just as it was capable of addressing tobacco policy. As the Supreme Court majority would later summarize

---

⁶⁶ § 103(g) encourages EPA research into the effects of CO₂, and § 602(e) requires that the EPA determine the “global warming potential” of various chemicals addressed in Title VI, which generally aims to protect the stratospheric ozone layer—a quite distinct goal from combating global warming. The provision in Title VI explicitly states that it is not to be used as the basis for additional regulation.
(Massachusetts v. EPA, 2007, 512), “In essence, EPA concluded that climate change was so important that unless Congress spoke with exacting specificity, it could not have meant the Agency to address it”—any “literalistic” argument from the statutory text notwithstanding.

Next, the EPA insisted upon its own ability to interpret the CAA for itself and claimed that its judgments should receive deference. Even if one were to find that the CAA could cover CO$_2$, they explained, § 202(a)(1) provides discretionary authority to the Administrator rather than creating a mandatory duty through use of the phrase “in his judgment.” Because the Administrator has never exercised such judgment, and might find that regulation under the CAA was imprudent for any number of reasons, the petitioners’ attempt to force the agency to adopt the novel reading should be rejected (EPA 2003b, 52929-30). The agency further justified a wait-and-see posture through reiterating the presence of continued scientific uncertainty about the causes and consequences of global warming (52930-32).\textsuperscript{67} EPA was essentially arguing that if their old law was going to learn a new trick, they should be the ones to decide that.\textsuperscript{68}

Finally, the EPA concluded its notice with a rather peculiar apologia explaining how an absence of EPA action would not be equivalent to a lack of any administrative action targeted at climate change issues. Pointing to a number of speeches and voluntary initiatives advanced by President Bush, as well as continuing non-regulatory actions being taken by the federal government, the agency seems to have been hoping to reassure those onlookers most worried about global warming that adequate tools already existed to meaningfully address the problem (52932-34). This is telling—rather than simply insisting that the problem was a legislative one,

\textsuperscript{67} It should be noted that, in making the case for patience, the EPA did not take the hardest line against applying the CAA to global warming possible. That is, the agency did not argue that there could be no changes in the scientific understanding of climate change sufficient to justify applying the CAA to the issue, only that such an understanding did not currently exist.

\textsuperscript{68} Of course, for those who believed the agency was in dereliction of its already existing statutory duties and that the public was suffering as a result, this argument would seem arrogant and unconvincing.
capable of being addressed only through congressional action, the EPA is framing the problem here as susceptible to a change in enforcement strategies given already existing powers. If its arguments about congressional actions regarding global warming were found unconvincing, judges might satisfy themselves with knowing that some kinds of executive action on the issue were underway and thus reject the interpretive framing of the problem *vis a vis* the CAA that the petitioners were offering.

**C. Judicial Reactions**

The D.C. Circuit was the first to review EPA’s denial of the petition, and the three judges fractured on the question of whether the matter was properly before the court in their 2005 opinions. Under the CAA, the D.C. Circuit has initial jurisdiction over challenges to all nationally applicable “final actions” under the CAA, while the district courts have initial jurisdiction over action-forcing petitions. The D.C. Circuit panel was divided about whether the matter was properly in front of them. Judge Sentelle was of the opinion that it was not, and would have dismissed the petitions for that reason, but Judges Randolph and Tatel agreed that it was appropriate to treat the agency’s decision not to regulate as final action subject to D.C. Circuit review—although they disagreed on the merits, with Judge Randolph basically accepting the EPA’s arguments about deference and Tatel the petitioners’ arguments about the CAA’s requirements. As a result, the three judges produced three separate opinions, with Judge Randolph’s serving as the majority because of Sentelle’s concurrence in the denial of the petitions, which he justified as reaching the consequence most similar to his own judgment of dismissal.
Given the importance of the issue and the lack of agreement among the D.C. Circuit judges, it was perhaps little surprise when the Supreme Court granted cert in 2006. The Supreme Court’s decision in April 2007 would also prove fractured on both standing and the merits—but their split at least had the property of producing a tidy majority. The split was nearly identical to that in Brown & Williamson: Justice Stevens wrote for the dissenters in that case plus Justice Kennedy, while this time the court’s conservatives (Scalia, Thomas, and now Roberts and Alito) dissented with two opinions: Justice Roberts writing that the plaintiffs’ standing should never have been recognized, and Justice Scalia writing that even given standing, the plaintiffs should have lost on the merits. Let us examine each opinion, with an eye to understanding why “old law, new trick” prevailed in the case of regulating greenhouse gases under the CAA where it had fallen short in regulating tobacco under the FDCA.69

The majority’s opinion first seeks to justify the petitioners’ standing. Oversimplifying, since the issue is less directly relevant to my arguments, Justice Stevens emphasized several factors. First, § 307(b)(1) of the CAA creates a rather liberal standing regime for challenging administrative decisions made under the CAA. Second, he found that the lead petitioner, Massachusetts, was “entitled to special solicitude in our standing analysis” because of its “stake in protecting its quasi-sovereign interests” (520).70 Especially given the benefit of the doubt that this unusual status conferred, Stevens found that Massachusetts was able to establish how an

---

69 An obvious thought here, which goes a long way toward deflating this chapter’s pretensions, is that the most salient thing we can say changed from Brown & Williamson to Massachusetts v. EPA is that Justice Kennedy changed his vote. There is clearly a large grain of truth to this observation, but since Justice Kennedy did not author any of the opinions in either case, I do not hazard any guesses about what exactly his thinking was. Rather, I provide what I believe is the best way of understanding a principled reason for treating the two cases differently by emphasizing the institutional contexts of the two cases. Contra assertions that the different outcomes in the two cases cannot be reconciled in any principled way (e.g., Moncrieff 2008, 595), my objective is to provide such a coherent explanation.

70 This “states are special” standing doctrine, supported largely by a citation to a century-old precedent that none of the plaintiffs’ briefs had thought to include, has elicited a great deal of commentary and criticism (e.g., Nolette 2011, 362-7; Stevenson 2007; Adler 2008; Mank 2009).
EPA refusal to regulate would contribute to harm that was actual, imminent, and capable of redress. The Commonwealth did this by discussing how the rise in sea levels caused by global warming would erode the value of its coastal lands, and then following up this point with an assertion that EPA inaction would exacerbate these harms (521-26).

Moving to the merits, the majority’s opinion closely follows the dissent from *Brown & Williamson* in its defense of reading statutes broadly. Once again, against claims that the old law in question had not been intended to deal with the new trick now contemplated, the Court’s liberals insisted that it was Congress’s prerogative to create flexible statutory instruments that could be adapted to changing information.

While the Congresses that drafted § 202(a)(1) might not have appreciated the possibility that burning fossil fuels could lead to global warming, they did understand that without regulatory flexibility, changing circumstances and scientific developments would soon render the Clean Air Act obsolete. The broad language of § 202(a)(1) reflects an intentional effort to confer the flexibility necessary to forestall such obsolescence. (2007, 532)

The CAA, like the FDCA, could be read as creating a broad charter for its administering agency, making a lack of specific intentionality on the part of the enacting legislators quite irrelevant.

In a different and revealing way, though, Justice Stevens’ majority opinion resembles the *Brown & Williamson* majority opinion: it finds that the statutory language is clear enough to justify finding against the agency at *Chevron* Step One. Like Justice O’Connor in the earlier case, Justice Stevens is quite insistent about just how decisive the statutory text is, saying that the court has “little trouble” finding that “[t]he statutory text forecloses EPA’s reading” (528). The definition of “air pollutant” found in § 302(g), he explained, “unambiguous[ly]” “embraces all airborne compounds of whatever stripe,” leaving no valid room for doubting that CO₂ should come under the statute’s regulatory power (529). Given this finding of clear applicability, the majority argues that the statute itself refutes the EPA’s further argument that it retains discretion not to regulate; the reliance in § 202(a)(1) to administrative “judgment” must be limited to the
question of whether a threat is posed, rather than giving the administrator “a roving license to ignore the statutory text” (533). 71

The majority also comes face to face with the Brown & Williamson majority and its extensive discussion of post-enactment legislative history. Stevens adamantly rejects the EPA’s argument that, just as many congressional enactments had regulated tobacco and thus rendered an “old law, new trick” interpretation of the older FDCA problematic, so too had congressional enactments after the passage of the CAA shown that it was never intended to address global warming. These other attempts to promote “collaboration and research” could only be understood as “complements” to strong regulatory action, Stevens argued, not conflicting substitutes (530). The “unbroken series of congressional enactments” that arguably implicitly prohibited the FDA from exercising jurisdiction in Brown & Williamson had no real parallel in the case of GHG regulation (see Baird 2004).

The two dissents (each joined by all four dissenters) take issue with the majority’s conclusions at nearly every juncture. Chief Justice Roberts’ dissent closely followed the precedent of Lujan v. Defenders of Wildlife (1992), and would have found that the plaintiffs (to whom the dissent would not have given “special” status) failed to meet the requirements of injury in fact, causation, and redressability under Article III (2007, 535-47). On the question of redressability, Roberts pointed to the mismatch of using the locally-oriented CAA to address a global problem, slamming the Court for blithely assuming that EPA regulations would “likely prevent the loss of Massachusetts coastal land” (546-49).

71 Note that by being so entirely dismissive of any possibility of statutory ambiguity, the majority could entirely ignore the idea that the statute would need to be absolutely explicit to justify undertaking such an economically consequential program of regulation. The majority thus failed to endorse the reading of Brown & Williamson that would have boiled its precedential value down to “in the case of very important policy matters, err on the side of executive branch restraint.”
Scalia’s dissent on the merits expressed disbelief in the court’s ability to find so much clarity in the statute that the EPA’s judgment should be rejected as contrary to the law. First, Scalia insists that nothing in the statute would compel the Administrator to make a judgment simply because a petition had been filed requesting he do so. Instead, he would accept as compelling the reasons offered by the EPA why prudence dictated delaying a final judgment as the other policy responses they described proceeded (549-54). Scalia then excoriates the majority for its conclusion that the CAA’s definition of “air pollutant” in § 302(g) necessarily includes CO$_2$ and other GHGs. He points out that the definition includes a substance only if it is an “air pollution agent”—a term which the act does not define. And given this crucial ambiguity, the agency’s reasonable interpretation—which states that there is ongoing uncertainty as to whether GHGs really do constitute harmful pollution akin to lung-choking industrial emissions—ought to be given Chevron deference (556-60).

D. Responding to the Supreme Court’s Ruling

The majority’s holding did not explicitly decide any regulatory questions for the EPA—rather, the Court required that EPA must directly confront the issue of regulating GHGs under § 202 and “ground its reasons for action or inaction in the statute” (534-35). The Bush administration began moving toward an endangerment finding for § 202, publishing an Advanced Notice of Proposed Rulemaking (ANPR) in July 2008, thereby setting the regulatory process in motion (EPA 2008b). This notice showed that the Bush administration was taking the Court’s demand seriously—it ran to 166 pages in the Federal Register. At the same time, it shows that the agency was moving rather tentatively in determining exactly what complying with

---

72 Scalia points out that the question of whether a substance is actually an “air pollution agent” is crucial to making sense of the definition, which otherwise would include “everything airborne, from Frisbees to flatulence” (558n2).
the Court’s mandate would entail. The published notice includes letters from Susan Dudley, then Administrator of the Office of Information and Regulatory Affairs, as well as from the Secretaries of Agriculture, Commerce, Energy, and Transportation, all of whom expressed serious reservations about regulating GHGs under the CAA. The ANPR solicited comments from the public regarding ways of meeting their concerns while also satisfying the Court’s mandate. Faced with these difficulties, the Bush administration left office without further action.

When one begins to delve into the statutory details of the CAA, the embarrassments caused by having to regulate CO₂ as an “air pollutant” are serious enough to make the EPA’s delay understandable. The implications of making an endangerment finding under § 202 are legion. Once GHGs have been designated as an air pollutant under the definition in § 302(g), it seems nearly impossible to offer a principled justification for failing to regulate industrial emissions. Specifically, it seems that EPA should have to create new source performance standards (NSPS) under § 111(b)(1)(A), which requires regulation of new sources (meaning new or modified plants) in any category which “causes, or contributes significantly to, air pollution which may reasonably be anticipated to endanger public health or welfare.” Additionally, it seems hard to justify not listing GHGs as a “criteria” air pollutant under § 108(a)(1), which requires inclusion of every air pollutant whose presence in the ambient air “may reasonably be anticipated to endanger public health or welfare.” Such a listing would trigger a host of other statutory requirements. Most importantly, it would require the EPA to set both primary and secondary National Ambient Air Quality Standard (NAAQS) for GHGs under § 109, based on a judgment of what concentrations of GHGs in the atmosphere would be “requisite to protect the public health” (for primary) and “requisite to protect the public welfare from any known or anticipated adverse effects associated with the presence of such air pollutant in the ambient air.”
(for secondary). That standard having been set, it would trigger a requirement for states to devise state implementation plans (SIPs) capable of bringing their local concentrations of GHGs into compliance with the NAAQS.

This statutory design is intended to force state and local governments to take actions to mitigate localized issues with breathable air, and threatens to create an absurd and impossible set of requirements when applied to the global problem of rising GHG concentrations. If the EPA set the NAAQS at any concentration lower than the prevailing world-wide concentration of GHGs, then the CAA’s literal requirements would require states to adopt a series of increasingly draconian measures to reduce their GHG emissions—which is to say, to reduce their industrial activities and energy consumption toward zero (Reitze 2009). In addition, best-technology standards would be required for nearly every new industrial source under § 111, and related language in § 165 and the permitting requirements of Title V would extend this requirement to all “major” sources—defined as emitting more than 250 tons per year and 100 tons per year of the regulated pollutant, respectively for each section. For CO₂, this would include nearly 40,000 sources for § 165, and approximately six million (!) sources (including many residential and commercial buildings) for Title V. Such an inclusive scope of regulation would utterly swamp the EPA’s administrative capacity (Adler 2011, 12-16).

The EPA under President Obama, headed by Administrator Lisa Jackson, has done its best to work through these many problems as it proceeds toward functioning regulation, but at every stage its choices have been met with litigated challenges that have yet to be resolved. EPA proposed its § 202 endangerment finding in April (2009a) and finalized the rule in

---

73 Of course, if the standard was set above prevailing concentrations, that would imply that all states were in compliance, thus requiring no further action and thereby defeating the purpose of the standard-setting.
74 The following discussion is drawn in part from Wannier (2010), who provides full citations to the cases involved.
A gaggle of dissatisfied industrial firms and trade groups, joined by a number of sympathetic states, quickly submitted petitions for agency review, which the EPA denied (2010d), leading to litigation. The petitioners in this action challenge the adequacy of the science behind EPA’s endangerment finding—and probably have relatively little chance of succeeding, according to several observers (Adler 2011, 9; Wannier 2010, 3). Substantive regulation under § 202 proceeded closely on the heels of the Endangerment finding (EPA 2010b). Consolidated litigation against these rules is also proceeding in the D.C. Circuit, though the agency’s logic on this substance seems to closely track the Court’s majority opinion; the petitioners nevertheless argue that the rule is redundant with existing CAFE rules administered by the Department of Transportation.

EPA is perhaps on shakier legal ground in several other areas where it has moved forward with GHG regulation under the CAA. In the so-called “Timing Rule,” the EPA proposed to reinterpret several sections of the act, including § 165 and its “prevention of significant deterioration” requirement, such that CO₂ would be covered even if EPA does not issue a NAAQS for CO₂ under §§ 108 and 109 (EPA 2010a). Industry petitioners have challenged this rule, claiming that EPA has inappropriately concluded CO₂ is “regulated” for the purposes of the whole act as a result of its activities under § 202, and that an absurd system would result if EPA is allowed to use § 165 to control GHG emissions (Wannier 2010, 14). Finally, confronting those purported absurdities directly, the EPA issued the so-called “Tailoring Rule,” which seeks to tailor the CAA’s statutory requirements to the needs of sensible GHG regulation (EPA 2010c). To do this, the EPA forthrightly says that it will ignore certain statutory requirements and adjust statutory thresholds for applicability based on the doctrine that an

---

75 Apparently, Administrator Jackson believed that the Supreme Court’s mandate in Massachusetts v. EPA straightforwardly obligated the EPA to make the endangerment finding (Adler 2011, 8n27).
agency should avoid absurd results; the doctrine that agencies may act out of “administrative necessity”; and a general appeal to broad discretionary powers entrusted to the agency under the CAA, especially by its “necessary and proper”-clause-equivalent in § 301(a). Challenges to this rule emphasize that EPA’s actions do not simply massage some minor procedural points—they run directly counter to the Act’s explicit text. They argue that the agency’s legal improvisation in trying to use the CAA to regulate GHGs shows how fundamentally ill-suited it is to that task, and they ask that the agency’s rules be invalidated. These seem to be serious challenges, and one might wonder whether even judges sympathetic to the attempt to regulate GHGs under the CAA (and follow the mandate of Massachusetts v. EPA) will be able to accept the agency’s need to fudge in such a brazen manner.

On other fronts, the EPA under Jackson’s leadership has not been quick enough to satisfy environmentalists (and their state allies) hoping to see the promise of Massachusetts v. EPA realized. After suing the EPA to force action under § 111, various groups entered into settlements outlining timetables for EPA regulation of power plants and oil refineries. EPA has now missed agreed-upon deadlines for both power plants (Broder 2011) and refineries (Gardner 2011), which the agency insists is due to the complexity of the rules being formulated rather than because of any political considerations. Presumably, the difficulties of applying the CAA’s rules to CO₂ play a large part. Finally, the EPA is also likely to face challenges from the most vehement environmentalists, who insist that it should formally list GHGs as criterion pollutants under § 108, issue NAAQS under § 109, and finally require SIPs under § 110. Apparently, the Center for Biological Diversity is committed to this position, despite the many apparent difficulties it poses (Adler 2011, 21).
In the process of actually executing its “old law, new trick” with the CAA and greenhouse gases, it has been tough going for the EPA. All the while, its efforts have engendered significant enmity from industry and their allies on the political right, who have decried the new regulations as hindrances to an already struggling economy likely to produce little benefit (e.g., Hayward 2009). A divided Congress, meanwhile, has failed to produce any legislation resolving the developing difficulties in one way or another. While passing a cap-and-trade bill that would have created a comprehensive regulatory system for GHGs, complete with emission credit trading markets, was an early priority of the Obama administration, in the end it floundered in the Senate as other matters (especially healthcare) won out (CQ 2009, 10-3-10-7). Since the Republican takeover of the House of Representatives in the 2010 midterm elections, there has been a flurry of activity in that chamber designed to explicitly strip the EPA of any power to regulate GHGs, but though many of these actions have won House majorities, none has had much traction in the Democrat-controlled Senate.

E. Analysis

What does the successful effort to force the EPA to apply the CAA to global warming teach us about the dynamics of “old law, new trick”?

As discussed above, the impetus for advancing a novel interpretation came because of how various actors were situated institutionally rather than out of any belief that applying the old statute to the problem at hand was the first-best policy solution. Increasingly alarmed by the problem of global warming, environmentalists would have been quite happy to see Congress pass new, well-tailored legislation of some kind. Failing that, they were not content to settle for
inaction when existing statutory language presented them some leverage to force the
government’s hand by insisting the problem could be treated as an interpretive one.

Why did the Supreme Court ultimately accept this push for the “new trick,” where less
than a decade before they had denied an effort that had the added benefit of an agency’s
backing? Like the *Brown & Williamson* majority, the majority in *Massachusetts v. EPA* found
that the broadly worded regulatory statute was nevertheless clear enough to foreclose the
agency’s favored interpretation. As in the tobacco case, the dissent seems to have the stronger
case in arguing that the statute was genuinely ambiguous. Justice Scalia’s dissent powerfully
assaults the idea that the definitional language of § 302(g) of the CAA self-evidently includes
CO₂ and other GHGs. Like Justice O’Connor’s *Brown & Williamson* majority opinion, he can
accuse the other side of “literalism” in the face of a great deal of evidence from the structure of
the act pushing in the other direction. Once again, however, the awkwardness of fitting the
existing statute to the problem seems to be insufficient in itself to resolve the issue.

The clearest point of commonality between the two majorities is the analysis of the
institutional context in which the “new trick” was being advanced. The *Massachusetts v. EPA*
majority explicitly considered whether other congressional enactments made a change in
interpretation unreasonably disruptive, as it had in the case of tobacco regulation, and they
concluded that they did not. Where opponents of regulating tobacco under the FDCA could
point to a number of powerful regulations already governing the market in tobacco, the best that
opponents of applying the CAA to global warming could muster were a few scattered efforts to
promote research on the issue. While the tobacco regulations could certainly be criticized as

---

76 Undoubtedly, some reader is here thinking, “Only Justice Kennedy can know for sure.” Note 69, above, discusses
this objection.
inadequate, global warming regulations could plausibly be portrayed as non-existent, creating a vacuum that a novel interpretation of the CAA could potentially fill.

The existence of this vacuum is certainly contestable—and, indeed, both of the dissents do directly contest it, giving much greater weight to congressional and presidential deliberations that had not yet produced any actions. It is telling that the Court’s conservatives join this issue directly, rather than simply conceding the institutional point but still insisting upon the constitutional procedures of bicameralism and presentment. The opening paragraphs of Chief Justice Roberts’ dissent are especially relevant: he begins by conceding the potential gravity of global warming, but insists, contra the majority, that “[i]t is not a problem, however, that has escaped the attention of policymakers in the Executive and Legislative Branches of our Government, who continue to consider regulatory, legislative, and treaty-based means of addressing global climate change” (2007, 535). The argument that Congress was “on the job” is certainly not wholly implausible—between 1999 and 2007, Congress introduced more than 200 bills to regulate GHGs, though none were enacted (Reitze 2009, 1). Sustaining this line of argument is made more difficult, though, by EPA’s pointing to a whole host of unilateral executive actions being taken by President Bush as it defended its decision not to apply the CAA. Apparently, by late 2006 and early 2007, when the Supreme Court decided the case, five justices were prepared to accept the idea that “old law, new trick” was better than the alternative combination of smaller-scale executive branch action and legislative inaction.

---

77 Moncrieff (2008) also argues that Congress and the Executive branch were actively addressing the issue of global warming prior to the Court’s decision, and on that basis argues that the Court’s holding was misguided and disruptive. She ultimately reaches a conclusion similar to this chapter’s: “Perhaps the greatest challenge for a reincarnated noninterference rule is to develop a standard for distinguishing serious congressional deliberation from strategic congressional posturing” (642).

78 A useful contrast can be drawn with another even more ambitious attempt by environmentalists to regulate GHGs through an “old law, new trick” maneuver. As the CAA battle was being waged, the Center for Biological Diversity (and other environmental groups) argued that in order to protect polar bears from extinction, the Fish and Wildlife Service (FWS) (part of the Interior Department) must promulgate regulations under the Endangered Species Act.
How does the series of legal controversies sparked by this choice reflect on their decision? Opinions will vary, no doubt. Arguably, something is better than nothing.

Institutionally speaking, it is not clear that it is appropriate to hold the Supreme Court responsible for the burgeoning legal difficulties in trying to apply the act, since at every moment Congress has had the power to set things straight one way or another. Indeed, politically speaking, “Congress will have to do something now, or else burdensome CAA regulation will follow” was one of the most politically effective arguments used to support the Waxman-Markey cap-and-trade bill (H.R. 2454, 111th Congress) that received House approval in 2009 but failed to

(ESA) to protect their habitat from the ravages of global warming, which requires nothing less than complete regulation of GHG emissions. Although the FWS under President Bush would eventually agree to list the polar bear as a threatened species (EPA 2008a), it subsequently issued a rule (EPA 2008c) outlining protective measures to be taken that did not include regulation of GHG emissions. The Center for Biological Diversity sued to force such regulations after President Obama’s Interior Secretary, Ken Salazar, affirmed the previous administration’s choice (Tankersley 2009). The litigation led to a recent District of D.C. decision, In re Polar Bear Endangered Species Act Listing (2011), in which District Judge Emmet G. Sullivan found that the Interior Department’s interpretations of the ESA were permissible, granting summary judgment on those points. (The Judge did grant environmentalists a lesser victory by finding that the FWS erred by failing to prepare a full environmental impact statement under the National Environmental Protection Act.)

In this case, unlike that of the CAA, the statute at issue would neither give the agency guidance in how to control emissions nor set any clear limits on the scope of regulations permitted (or required) under its auspices. It is also impossible to draw a tight causal connection between a particular source’s emissions and identifiable damage to polar bear habitat. These factors alone were sufficient to doom the challenge to the FWS interpretation of the ESA, as Judge Sullivan’s opinion is confined entirely to a considering the propriety of the agency’s interpretation given its statutory requirements. Considering institutional competencies can also provide a strong reason for rejecting their challenge, though. The FWS has no expertise in regulating ambient levels of atmospheric compounds comparable to that of the EPA, but just as importantly this attempt to have an old law perform a new trick came in the shadow of the already successful “old law, new trick” campaign that culminated in Massachusetts v. EPA. Given the ongoing development of a regulatory system covering GHG emissions under the CAA, it would seem to be superfluous (and possibly quite confusing) to ask for a parallel regime of regulation to be built on the ESA.

238
win Senate approval.\textsuperscript{79} As of this writing nearly six years after \textit{Massachusetts v. EPA}, legislative recalibration has yet to arrive, though it is possible that it will yet.\textsuperscript{80}

**IV. Conclusion**

Although I have argued that a single coherent logic can support the holdings in \textit{Brown & Williamson} and \textit{Massachusetts v. EPA}, there is certainly the temptation to read the two cases’ divergent aftermaths as figures in a cautionary tale against accepting “old law, new trick” maneuvers. Although anti-tobacco activists were undoubtedly disappointed by the result in \textit{Brown & Williamson}, it enabled them to proceed without any illusions that existing statutes provided a sufficient statutory resource for addressing their issue. Although a durable compromise took a decade to hammer out, it was eventually forthcoming. Now that the FSPTCA is in place, the FDA is equipped with statutory tools much better suited to the purpose of tobacco regulation than those the FDCA could have provided. For the environmentalist winners of \textit{Massachusetts v. EPA}, though, achieving policy change without statutory change has been a decidedly mixed bag. Regulation has proceeded onerously, confronted by litigation at every turn. Many thousands of man-hours (perhaps millions?) are being devoted to devising an awkward application of the CAA apparatus to global warming. Whether there is eventually a bipartisan compromise or a Republican-sponsored-and-signed bill depriving the EPA of its

\textsuperscript{79} Title VIII, Part C of the bill would have removed EPA jurisdiction of GHG emissions from stationary sources, though the EPA could still regulate mobile source emissions. Phil Barnett, Representative Waxman’s top aide in the Energy and Commerce Committee, confirms that support for the bill was significantly bolstered by the sense that impending CAA regulation would be problematic (informal communication with author, 2012). Some have questioned whether the “threat” of CAA regulation would effectively force Congress to act; for example, Stavins (2010) asks “whether this is a credible threat, or will instead turn out to be counter-productive (when stories about the implementation of inflexible, high-cost regulatory approaches lend ammunition to the staunchest opponents of climate policy).”

\textsuperscript{80} Alternatively, one can imagine how the sense that “something is happening” could lessen the impetus for compromise on new action in Congress, but there seems to be little evidence that this effect has occurred in the case of greenhouse gasses.
jurisdiction, this monumental effort is likely to be rendered more or less meaningless (at least as policy, rather than as a source of political leverage). This must be a source of tremendous frustration for those laboring within the EPA and it certainly represents a deadweight loss for our society.

Even conceding this waste, though, the advocates of teaching the CAA this new trick are likely to say: what choice did we have? Faced with an ineffectual legislature and a dithering agency, they decided the prospect of action using existing statutory tools was appealing. It is easy to wonder if they made the right choice—but hard to say what the best alternative was for advocates of policy change, pragmatically speaking. Many environmentalists assert without hesitation that the CAA program fighting global warming is far superior to no program at all. When they think of the majority opinion in *Massachusetts v. EPA*, they see five justices muddling through by supporting the best available option, given the all-too-human Congress we have. They are likely to be quite dismissive of theoretical arguments that would insist that judicial actors should always conceive of themselves as agents of the Platonic ideal of the Congress we would want.

Keeping this perspective in mind, we should avoid over-reading the two cases and see that the right question to ask isn’t whether old laws should ever be taught new tricks, but when doing so is appropriate. However one is inclined to answer this question, judgments about institutional capacities are likely to be at the heart of the matter. These judgments are unlikely to displace considerations of statutory text, intent, or purpose—though given a dramatic enough circumstance, they might.81 Judges differ in the weight they give to each of these factors, but

---

81 If a court somehow came to hear a challenge to the Treasury Department’s use of the Exchange Stabilization Fund to stabilize money market funds (discussed in notes 10 and 20), the institutional superiority of the executive department over Congress in addressing an extremely grave and time-sensitive threat would undoubtedly have
few forsake any of them; when we talk about “textualists,” “purposivists,” and “pragmatists” among practitioners we are really talking about differences in emphasis rather than kind.

Apart from these different weightings, judges also have different visions about what a well-functioning policymaking system looks like. Conservatives (in the little-c sense and in the anti-regulatory sense) are more likely to believe that inaction on some “problem” represents an acceptable working of the Madisonian separation of powers and that status quo bias is a generally healthy thing. Those who are less certain about the ongoing fitness of our constitutional regime, as well as those inclined to view regulation as indispensable in the modern world, are likely to see inaction as evidence of dysfunction. While these disagreements may seem arcane, in fact they go directly to the nature of our federal government.

Merely discussing these issues explicitly is unlikely to lead to any particularly dramatic changes—after all, if I am right, thinking in the institutional vein is already widespread and probably well-understood by most practitioners. As I indicate throughout this dissertation, it could, however, help to open a new horizon for constructive empirical research. Judges are forced to make determinations about institutional competencies almost entirely based on their own “casual empiricism”—a phrase from the legal literature I don’t use disparagingly. This fact helps to explain how different judges can see things so differently, since each must rely on her own impressions. Scholars should aspire to provide systematic empirical accounts of policymaking patterns and consequences that could constructively settle some of these questions, leading to a more rational and consensus-based policymaking system.

weighed heavily on the judges’ minds, perhaps leading them to overlook the near non-existence of support for the action in statutory text, intent, or purpose.
Cases Cited

U.S. v. 23, More or Less, Articles, 192 F.2d 308 (2nd Cir. 1951)

62 Cases, More or Less, Each Containing Six Jars of Jam v. U.S., 340 U.S. 593, 600 (1951)


United States v. 46 Cartons, More or Less, Containing Fairfax Cigarettes, 113 F. Supp. 336 (D.N.J. 1953)


Action on Smoking and Health (ASH) v. Harris, 655 F.2d 236 (D.C. Cir. 1980)


U.S. v. 25 Cases, More or Less, of an Article of Device, 942 F.2d 1179 (7th Cir. 1991)


U.S. v. An Undetermined Number of Unlabeled Cases, 21 F.3d 1026 (10th Cir. 1994)

Coyne Beahm v. FDA, 966 F. Supp. 1374 (M.D.N.C. 1997)

Brown & Williamson Tobacco Corp. v. FDA, 153 F.3d 155 (4th Cir. 1998)


Massachusetts v. EPA, 415 F.3d 50 (D.C. Cir. 2005)


Regulatory Materials Cited

Environmental Protection Agency (EPA)


2003a. Memorandum by Robert Fabricant, EPA General Counsel. “EPA’s Authority to Impose Mandatory Controls to Address Global Climate Change under the Clean Air Act.” August 28.


2010d. “EPA’s Denial of the Petitions To Reconsider the Endangerment and Cause or Contribute Findings for Greenhouse Gases Under Section 202(a) of the Clean Air Act.” 75 Federal Register 49556, August 13.

Food and Drug Administration (FDA)

1995a. “Regulations Restricting the Sale and Distribution of Cigarettes and Smokeless Tobacco Products To Protect Children and Adolescents; Proposed Rule.” 60 Federal Register 41314, August 11.

1995b. “Analysis Regarding the Food and Drug Administration’s Jurisdiction Over Nicotine-Containing Cigarettes and Smokeless Tobacco Products; Notice; analysis regarding agency jurisdiction.” 60 Federal Register 41453, August 11.


**Reporting on Congress**


**Works Cited**


Chapter Five – Statutory Erosion and Maintenance: The Case of the Glass-Steagall Act

What does it take for a regulatory law to be effective over a span of decades? Instinctively, it is easy to believe that all that is necessary is benign neglect on the part of future legislators and simple compliance on the part of those charged with executing the law. If policy change is a function only of statutory change, an absence of relevant new legislation implies stability in government actions. The law endures, unchanging as a Platonic form.¹ As Chapter Four explains, however, changes in enforcement priority or in the interpretation of statutory language can lead to changes as well. These mechanisms of change create the possibility of “teaching old laws new tricks,” thereby expanding the reach of the policymakers charged with administering the statute.

In this chapter, I demonstrate how interpretive changes can work in the opposite direction, eroding a statute’s practical effects rather than broadening them. Regulated interests supply the motive force behind statutory erosion by discovering and exploiting loopholes in complex regulatory statutes, on which I focus here. Resisting this process requires statutory maintenance, accomplished either through targeted legislative amendments or interpretations making the statute’s practical operation consistent with enacting legislators’ goals. As in Chapter Four, we see that creative interpretations can serve as near-substitutes for legislative changes, but not total substitutes. Because some statutory text remains clearly binding, old statutes whose practical effects have eroded nevertheless shape and constrain government actors

¹ Many of our models for thinking about the policymaking system do treat law more or less in this way, as shown in the Introduction to this dissertation. In their seminal article, Landes and Posner (1975) begin from the peculiar assumption that judges will make law work in this manner because doing so will allow legislators to produce more durable laws and thereby maximize their leverage in bargaining with interest groups.
as they attempt to cope with changing circumstances, sometimes in ways that no one would desire.

Policy erosion can be illuminated by analogizing the enactment of a public policy to building a dike, with the goal of the policy represented by the goal of keeping all the water on one side of the dike. Without attention, small holes may develop into gaping fissures. Some can be plugged by a timely finger; others may require more serious reconstruction. This essay’s focus on the history of the Glass-Steagall Act, a law universally described as having erected a wall between commercial and investment banking activities, makes the metaphor of the dike particularly apt. In general, the metaphor offers a useful way of thinking about systemic regulations, which require across-the-board structural integrity in order to be effective but are complex enough to make maintaining this difficult.²

While the concept of erosion is often employed casually when talking about regulation (public figures discussed in the case study often employed the metaphor themselves), no systematic work on the subject exists. The idea resonates with contemporary scholars’ theory-building, however. Hacker (2004) discusses ways in which welfare state opponents have avoided frontal attacks upon venerable statutes and instead sought to prevent needed updating, leading to “policy drift” in which on-the-ground actions no longer match statutory goals.

² More generally the metaphor will not be applicable to all types of laws. Some laws have clearly delineated scopes, leaving little potential for complex legal arguments to erode the intended policy effect. Failures in execution may still prevent the achievement of policy goals, but this will not be a problem of legal interpretation. On the other hand, there are some policies crafted so as to ensure not only their own maintenance, but also the growth in their absolute and relative importance over time. For example, the incentives created by the structure of Medicaid virtually guaranteed that the program would ratchet ever-upward in scope and cost (Greve and Wallach, 2008). No maintenance would be necessary to secure these results. For these types of policies, a snowball rolling down a snowy hill and gathering momentum may a more appropriate metaphor.

The question of what characteristics push a policy toward erosion, stasis, or growth is an extremely important one, and many works on the creation of bureaucracies already consider the ways in which agencies are built to reflect their enacting coalitions (Moe 1989; Lewis 2003), or put agencies on “autopilot” and thus ensure that they reach certain end-goals (McCubbins, Noll, and Weingast 1989). This paper mostly takes for granted the existence of eroding forces, but the relevance of erosion, maintenance, and momentum to decisions of how to structure legislation is a subject that would reward further exploration.
Patashnik (2008) gives an account of political erosion of reform laws, and the necessary conditions for maintenance. He shows that reforms enacted on behalf of a diffuse public and opposed by well-organized interest groups, which are difficult to pass in the first place (Arnold 1990), are likely to be undermined even by those who managed to strike a bargain for the public good in the first place. Such reforms are vulnerable to a fate of “death by a thousand cuts” in the political process, and will only be capable of enduring if they attract backers devoted to fending off eroding amendments. This essay will also emphasize the importance of a law’s having strong political supporters capable of doing maintenance, but makes a unique contribution to this literature by focusing on the operation of erosion and maintenance through legal processes, thus building on works that feature bureaucrats and judges shaping policy development through legal interpretations (e.g. Derthick 1979; Melnick 1983, 1994). There is also an extensive literature that thinks about the ways in which bureaucrats and especially presidents (whose role perhaps receives short shrift in this chapter and dissertation) use “administrative” methods and “unilateral” actions to shape policy rather than seeking change through the legislative process (e.g., Moe 1985; Durant 1987, 2009; Howell 2003; for a survey of recent Presidency research, see Moe 2009). In many cases, these works discuss what might be called “erosion” through non-interpretive mechanisms—largely what Chapter Four called questions of enforcement. While clearly examining very similar phenomena, the authors writing in this tradition tend to overlook the way statutes guide and constrain executive branch actors, as well as the distinctively legal dimensions of their administrative actions. Lists of executive branch means of control, which generally include departmental reorganization, control over budgets and personnel, appointment of loyal political executives, and controlling enforcement priorities (Durant 1987, 180) often neglect control over agency statutory interpretations. This dissertation takes a small step toward
bridging the gap between this public administration literature and the statutory interpretation literature, although it fails to address important questions about presidents’ motivations and the important role of the Office of Management and Budget in reviewing policies emanating from statutes as they are interpreted.

Exploring the dynamics of erosion also provides new and important insights into regulatory capture, also known as the “interest group theory of regulation” (Stigler 1971; Peltzman 1976). Studying erosion lets us understand one of the central mechanisms of capture rather than viewing it as the result of a black box process. The perspective I offer allows us to see capture as a mundane part of a larger process of legal contestation rather than as a threatening and mysterious force in our policymaking process. Bureaucrats’ actions are best understood as part of muddling through, under law, in an attempt to balance competing substantive goals, fidelity to statutory text, legal coherence, and programmatic workability. Thinking in these terms about what is necessary to maintain a policy’s efficacy, meanwhile, allows critics of capture a way to think concretely about means of resistance—as well as realistically understanding limiting factors and likely causes of failure.

After elaborating on the concepts of erosion and maintenance and their relationship to statutory interpretation, this paper offers an in-depth case study of the development of the Glass-Steagall Act, which was maintained for many years but eventually gave way to erosion after bureaucrats’ priorities shifted.3 Deep engagement with the case study elucidates the dynamics of maintenance and erosion—as well as shedding light on a policy of great interest in light of claims that the demise of Glass-Steagall played an important role in exacerbating the financial

---

3 This article is not the first to notice the general dynamic described here. In particular, Kaufman and Mote (1990) offer an excellent descriptive account through 1990 that provided a helpful starting point for the current inquiry.
crisis of 2008. As always, however, deriving insights through a close study of one policy area leaves for future work the question of how similar other areas are.

**Erosion, Maintenance, and Statutory Interpretation**

Legislators enacting a new regulatory scheme pursue their policy goals largely by setting down legislative text directing the course of bureaucratic action. The specific commands in the regulatory scheme leave room for executive branch interpretation, through clauses explicitly delegating choices to administrative judgment or by defining key terms incompletely and leaving statutory gaps that bureaucrats are obliged to fill in as they apply the law to concrete facts. In a complicated, dynamic world peopled by creative lawyers who probe every point in the statutory scheme for potential loopholes, certain interpretive choices are necessary to align the statute’s practical operation with enacting legislators’ expectations. I call choices that diminish a statute’s ability to accomplish its original goals *erosion*, and choices that enhance its efficacy *maintenance*. As the narrative below will show, maintenance may enhance a law’s efficacy either through directly promoting its underlying purpose or by ensuring that the policies it creates remain coherent, clear, and capable of enforcement.

Depending on the original statutory language and the ingenuity of erosion-seekers, many perfectly plausible readings of the law may lead to erosion of important goals; on the other hand, some parts of the law may leave little room for interpretation, clearly demanding certain actions on the part of bureaucrats, even if they no longer prioritize pursuing the goals embodied there and would favor erosion as a matter of policy.

In most regulatory regimes in the modern American administrative state, interpretive choices are made in the first instance in the executive branch by a combination of career
bureaucrats, lawyers, and political appointees (Strauss 1990, Mashaw 2005). Executive interpreters’ choices are made in the shadow of potential judicial intervention, presuming that the judiciary has litigants to present it with cases over which it has jurisdiction.

Because of the possibility of judicial review—and also because they want to get the law “right” for its own sake—bureaucrats think about the same interpretive desiderata as described in Chapter Four, including the content of the statutory text (Section I.A) and the intent and purpose of the statute (Section I.B). Depending on how the statute is written and what kind of pressures it faces, erosion or maintenance could come about through quite different interpretive strategies, both by bureaucrats or by judges.

Textualism facilitates erosion if the statute is somehow self-defeating, either because it contains loopholes that can be exploited, its scope excludes relevant actors, or the remedies available to enforcers are insufficient to give the law bite. This may frustrate those committed to what they see as they law’s underlying purpose, but, as we saw in Chapter Four and will soon see again in this chapter’s case study, a commitment to clear text’s ability to bind seems to require that maintenance only be undertaken by the people’s elected representatives in Congress, with bureaucrats and judges constrained to dealing with the law as it is. On the other hand, if the statute is written such that its clear application ensures regulatory stringency, then the textualist approach can constrain attempts to promote regulatory flexibility by working around vexing

---

4 Where do loopholes come from? Many are purposefully included in statutes in the hopes that creating a narrow exemption to a general scheme can advance an important policy interest without jeopardizing the overall effectiveness of the statute. These intentional loopholes can be thought of as counterparts of “red tape,” which Kaufman (1977) characterized as minor regulations enacted to ensure various forms of fairness in bureaucratic action, each independently enacted with good intentions but without consideration for their cumulative effects. Just as red tape is rarely meant to be burdensome, few loopholes are designed to undermine the overall functioning of the statute in question. It is perhaps an iron law of regulation that more private actors will take advantage of loopholes than their creators anticipated. Unintentional loopholes are just as important; these result from the imprecise drafting of statutes caused by the limited foresight of those framing rules.
provisions. In this case, the textualist response recalls Mencken’s definition of democracy\(^5\): the people should get what they asked for, through their legislators, whether executive branch policymakers judge their choices prudent or not. In this case, maintenance requires adherence to the text’s clear requirements in the face of pressures to derive policy change through a new interpretation.

When text is less than completely clear, purposive interpretations of indeterminate statutory text are frequently used to support statutory maintenance. If treating a statute’s text narrowly in the name of fidelity will impair its ability to fulfill its underlying purpose (as the interpreter sees it), those fighting against erosion will make purposive arguments about remaining faithful to the law’s true spirit. Can purpose-based arguments conversely lead to erosion? This seems harder to fathom; after all, though a statute’s text might defeat its true purpose, it is hard to see how the purpose could defeat its purpose. It is important to remember, though, that legal interpreters do not take up one statute at a time, but instead are forced to interpret and put into effect the whole body of statutory law all at once. In this larger enterprise, particular statutory requirements are bound to be in tension. Statutory purposes, then, are plural, and pursuit of one may mean neglect of another. Bureaucrats will justify their interpretations as balancing multiple, sometimes-conflicting goals, requiring expert judgment and an intimate knowledge of the tradeoffs embedded in specific regulatory decisions—characteristics judges typically lack. Under the doctrine of *Chevron v. NRDC* (1984), they will therefore ask for deference (see Chapter Four, Section I.C) from judges as they muddle through, (professedly) under law. Appeals to statutory purposes in making sense of ambiguous statutory language, then, are likely to cut in favor of bureaucrats, who will almost certainly be able to offer credible

---

\(^5\) “Democracy is the theory that the common people know what they want, and deserve to get it good and hard” (Mencken 1916, 19).
arguments for why their daily engagement with the law and its applications make them, and not judges, best suited to determine what that law should mean in practice.⁶

Precisely because so many purpose-based arguments are available to interpreters, many have warned that if purposive arguments are accepted as justifications of text-stretching interpretations then nothing is effectively prohibited to executive branch interpreters. Citing public choice scholarship demonstrating the manipulability of majority voting procedures, such critics question whether the idea of a collective purpose or intent is even at bottom a meaningful one (Farber and Frickey 1988; Shepsle 1992; McCubbins, Noll, and Weingast 1994). Strong textualists drawing on this tradition argue that if a determinate meaning can be extracted from the text, that should be the end of the story (see Vermeule 2005, 2006; the discussion in the Introduction regarding Cell IV.D; and the discussion in Chapter Two, Section 5.B).

As previous chapters have discussed at length, this logic is persuasive as far as it goes but begs two questions: how clear must text be to be treated as completely determinate, and what are interpreters to do if it seems to them that their fundamental goals will be thwarted by doing what statutory text seems to command? Difficult interpretive disputes arise precisely because these questions resist any once-and-for-all resolution. Instead, both bureaucrats making their interpretive choices and judges weighing the permissibility these interpretations face particular sets of institutional circumstances that profoundly shape their decisions. There is a constant balancing act between the two aspects of the rule of law. On the one hand, legal interpretations must all hang together to produce coherence and workability in the system as a whole. On the other hand, obedience to text must count for something if laws are to be worth the paper they are printed on. Bureaucrats testing the limits of statutory text, whether in service of erosion or

⁶ As Chapter Four makes clear, though, if the interpretive question at issue is itself about the nature and scope of the functions a statute is to have, this logic will provide lesser support for deference.
maintenance, will stress the former; interests disadvantaged by their choices stress the latter, and say that Congress must be the one to change the statute’s requirements.

For this textualist argument to carry weight, judges must believe that Congress is willing and able to pursue meaningful action in the relevant policy areas, such that its failure to act conveys meaningful information about a preference for preserving the status quo. As the case of Glass-Steagall will demonstrate, however, this belief in legislative ability can only withstand so much falsification before it gives way to the inclination toward deference. As in Chapter Four, context-specific institutional capacities will guide judges’ thinking about how much leeway to give executive branch creativity. The account of Glass-Steagall’s bureaucrat-driven erosion, and the courts’ ultimate acquiescence to it in the 1980s, suggests that judges are likely to allow even text that seems quite fixed in its meaning to be bent if the institutional facts-on-the-ground suggest that political support for legal change exists but is unable to mobilize legislative action. “Clarity” begins to look less like a property of text than an indicator of existing political will. Because some text is simply too clear to ignore, however, a largely vitiated statute can continue to constrain policy in distortionary ways.

Table 1 schematically organizes the preceding discussion, and will be helpful in understanding the maintenance and erosion of Glass-Steagall.
Table 1 – Situations in which Textualism and Purposivism Promote Erosion and Maintenance

<table>
<thead>
<tr>
<th></th>
<th>Textualism</th>
<th>Purposivism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erosion</td>
<td><strong>If</strong> Statutory text naturally self-defeating (loophole, scope too narrow, remedy provided inadequate, etc.)</td>
<td>Goals (internal or external) in conflict, so that pursuing one naturally impairs another</td>
</tr>
<tr>
<td></td>
<td><strong>Then</strong> Let it defeat itself. “Congress’s problem to solve by fixing law”</td>
<td>Defer to bureaucrats’ balancing of goals, accepting result as the product of expertise and care</td>
</tr>
<tr>
<td></td>
<td><strong>When</strong> Instinctive for judges, but hard to keep up if Congress unresponsive; hard to accept for bureaucrats unless they are pro-erosion</td>
<td>Natural part of bureaucracy’s responsibility for multiple complex regulations; good faith arguments by executive likely to win judicial deference if they have some textual support</td>
</tr>
<tr>
<td>Maintenance</td>
<td><strong>If</strong> Unstrained interpretations of plain text will effectuate clear goals of statute</td>
<td>Statutory text naturally self-defeating, or clever privileging of form over function threatens to undermine fundamental goals</td>
</tr>
<tr>
<td></td>
<td><strong>Then</strong> Insist on straightforward interpretation when more creative reading would nullify intended force of law (The “textbook” case)</td>
<td>Give force to spirit of law even if letter lacking; reject narrow readings that vitiate law’s force, even if facially plausible</td>
</tr>
<tr>
<td></td>
<td><strong>When</strong> Perhaps the norm, but limited; unlikely to persist in changing world without periodic congressional interventions</td>
<td>Tempting for bureaucrats and judges who believe in regulatory goals; rarely openly admitted to; perhaps limited in effectiveness</td>
</tr>
</tbody>
</table>

Analytical Narrative: The Career of the Glass-Steagall Act

A. Origins

The Banking Act of 1933 (Pub. L. No. 73-66) was an ambitious and complex statute that regulated or prohibited practices believed to have brought on the Great Depression. Thanks to the policy entrepreneurship of Representative Henry Steagall (D – AL), the act’s most lasting contribution was to create national deposit insurance, an idea that had long been disfavored and still lacked the enthusiasm of leading politicians, including President Roosevelt (Schroedel 1994,
chap. 3; Kennedy 1973). Sections 16, 20, 21, and 32, which became known as the Glass-Steagall Act, were primarily the work of the architect of the Federal Reserve System, Senator Carter Glass (D – VA). These sections legally mandated the separation of commercial and investment banking, which until then had often been conducted by the same powerful firms.\(^7\) The law would purportedly prevent: 1) Banks making risky investments in securities, thereby endangering deposits; 2) Unsound loans made to prop up companies in which a bank was invested; 3) Pushing underwritten securities onto banking customers (Bentson 1990, 11-12).

Whether the law was a well-targeted intervention remains controversial. Congressional hearings conducted from 1931 to 1933, especially the Pecora Hearings in 1933, offered extensive evidence of the abuses mentioned above, but some historians have offered rebuttals, arguing that Pecora’s claims were anecdotal, misleading, and false (Bentson 1990). Whatever the truth, the regulations satisfied widespread demand for action in the wake of the financial system’s cataclysmic meltdown. In Kingdon’s (1995) terms, banking was catapulted onto the agenda by the crisis, and there was a paucity of policy alternatives available to answer the political needs of the moment. Perhaps just as importantly, Glass’s ideas were also highly congenial to dedicated securities firms, which would greatly benefit from having most of their competition banned. In 1933, the wider banking industry readily acquiesced to the regulations, as many primarily commercial banks with troubled balance sheets were contracting their operations and would not have been conducting securities activities in any case. Bankers were also well aware of their political situation, and may have thought that bowing to this reform was a relatively painless way to avoid more retributive measures (Bentson 1990, chap. 11). Until the 1950s, most banks were satisfied to maintain symbiotic relationships with separated investment houses.

\(^7\) Commercial banks accept deposits and make loans, while investment banks underwrite and market securities as well as executing various trades. When the term “bank” is used unmodified, this refers to a commercial bank.
B. Early Implementation, Erosion, and Maintenance: 1933-1956

Preventing erosion of Glass-Steagall’s “wall of separation” between commercial and investment banking was required soon after the law’s passage. The first two decades after its passage featured statutory maintenance by judges, bureaucrats, and legislators, all of whom sought to define the act’s practical applications in sustainable terms that prevented easy evasion.

One early challenge to the law’s efficacy was rooted in its central language, which prohibited anyone “primarily engaged” in securities activities from also engaging in commercial banking. In *Board of Governors v. Agnew* (1946), the Supreme Court was asked to review the Federal Reserve Board’s decision to remove as directors of a national bank several men who were indisputably involved in securities activities. First, the Court had to decide whether Federal Reserve decisions made under Glass-Steagall were subject to judicial review, which the Board of Governors disputed. For a unanimous Court, Justice Douglas decided this momentous question in the affirmative without offering much justification (444), ensuring that bureaucratic decisions about Glass-Steagall would receive judicial scrutiny. Second, the Court had to decide whether “primarily engaged” would be given a broad or narrow meaning. As the Fed read it, “primarily” meant “substantial,” whereas the removed directors argued it should mean “principal,” i.e. “greater than fifty percent” (446). The Court ruled that the Board’s broad

---

8 A note on the various banking regulators involved is in order. Banking has been subject to federal regulation longer than any other part of the economy, and over time it has acquired a multitude of separate regulators, each of which has its own distinct “personality” in addition to having different responsibilities. This means that each pursued its own unique agenda as it enforced Glass-Steagall. The Office of the Comptroller of the Currency (OCC) is responsible for chartering and regulating national banks and has tended to craft regulatory solutions in order to satisfy its “customers”— and since bankers seeking a new charter are faced with a choice of either national or state regulation, seeing the regulated interests as customers or clients is somewhat more apt for banking than other areas of regulation. This culture tends to overwhelm any effects of partisanship, and Comptrollers of both parties tend to be highly sympathetic to commercial banking interests. The Federal Deposit Insurance Corporation (FDIC) is built and acts like an insurance company but is also charged with regulating state banks that choose not to be members of the Federal Reserve System. When its insurance function was not directly implicated, it tended to follow the lead of the other regulators on Glass-Steagall issues. The Federal Reserve Board of Governors (Fed) is responsible for ensuring the safety and soundness of its member banks, and the BHC Act of 1956 also vested it with regulatory responsibility for Bank Holding Companies. The Fed functions somewhat like an academic institution, and generally acts as a repository of prevailing technocratic impulses regarding banking (Khademian 1996).
reading was both reasonable and fully consonant with the purposes of the statute, and therefore
due deference (447).

“Maintenance” may be an awkward label for this early case, since an edifice of active
enforcement had yet to be built. Nevertheless, it is easy to see how the Court could have thrown
its weight behind erosion by favoring a different interpretation of the statute’s plain language.
Given an ambiguous statute and bureaucrats invested in achieving the law’s underlying purposes
through purposivist interpretations, deferential court rulings such as this supported meaningful
implementation of Glass-Steagall’s wall between commercial and investment banking.

An important eroding development in the early years of Glass-Steagall was the migration
of many bankers to the bank holding company (BHC) corporate form, in large part because it
was not covered by the main provisions of Glass-Steagall or indeed by much of any regulatory
law. By bringing bank ownership under the same corporate umbrella as companies pursuing
non-banking activities, a BHC could effectively join control of banking and non-banking
activities, thus undermining Glass-Steagall’s purpose. Congress was well aware of this source of
potential erosion of the separation between commercial and investment banking, and bills
addressing it were introduced into almost every session of Congress after 1938, when Franklin
Roosevelt specifically called for regulation of BHCs. Major efforts in 1938 (Glass-McAdoo),
1947 (offered by Fed Chairman Mariner Eccles), and 1949-50 (Maybank-Robertson) died in
committee, opposed by the American Bankers Association (ABA) and its sympathetic regulator,
the OCC. By the early 1950s, the Federal Reserve was sounding a consistent refrain to pass

A Bank Holding Company is defined, rather unhelpfully, as any company exercising control over a bank. The
form offers various advantages for raising capital through issuing stock or debt. The Glass-Steagall Act put BHCs
under weak Federal Reserve Supervision, and never covered BHCs with only one bank (Jessee and Seelig 1977, 8).
Figure 1, appended to the text, clarifies the rather byzantine relationship between BHCs, banks, bank subsidiaries,
and bank holding company affiliates. It should be noted that putting a bank under the auspices of a BHC would not
relieve it from supervision by its normal regulator, but by creating an organization outside of that regulator’s
jurisdiction those people effectively exercising supervisory control might nevertheless evade regulators.
strong regulation of all BHCs (Jessee and Seelig 1977, 9). A critical source of disagreement that doomed these bills in committee was the question of whether regulation should extend to BHCs controlling just one bank (Worsham 1996, 90-99). With President Eisenhower’s support, a bill finally emerged in 1956 and Congress passed the BHC Act (12 U.S.C. § 1841, et seq.) by an overwhelming margin, establishing Federal Reserve regulation and prohibiting BHCs from managing or controlling non-banking assets other than those directly related to their core banking business. This was only partial maintenance, however: in order to win easy passage, the act included a slew of exemptions, most crucially excluding BHCs controlling only one bank, a source of disappointment for the president and the Federal Reserve. At time of passage, only 46 BHCs were subjected to regulation, while 116 (mostly very small) BHCs were exempted (CQ 1956, 557).

10 Consequential as it was, BHC Act supporters secured a consensus only by retaining many exemptions, demonstrating how legislation intended to promote the general good is often greased with side payments to affected interests in order to secure passage, thus undermining the very rationale that motivated the reform (see Patashnik 2008, 171). The price of this consensus was relative weakness, and by 1958 the Fed would be lobbying Congress again to close the act’s loopholes, especially the one-bank exemption (Jessee and Seelig 1977, 12).

C. Maintaining Separation under the BHC Act, 1956-1971

The next period of Glass-Steagall (and now BHC) development was characterized by the same dynamic of seeking to find ways to perform maintenance at the least possible political cost.

---

10 Worsham (1997) argues that the net effect of the BHC Act of 1956 was to facilitate erosion. His interpretation is that the New Deal coalition that had passed Glass-Steagall in 1933 was no longer energized to protect its own regulatory regime for banking, and therefore was not prepared to expend political capital to secure true, stringent maintenance. While maintenance certainly could have been more complete, his argument is misleading in two ways. First, to decide whether the BHC Act itself represented erosion or maintenance we need to use the status quo as our point of comparison, not the hypothetical provision of complete maintenance. Second, it is doubtful that Congress included the exemptions to undermine the whole system, and indeed they closed these loopholes in the following decade once they became understood as serious threats.
Bureaucrats and legislators remained basically committed to the original aims of Glass-Steagall, but they nevertheless sought to accommodate various constituents as they reacted to new developments. Where this accommodation seemed to go against Glass-Steagall’s purpose, the courts intervened to perform maintenance of their own.

The major challenges of erosion during this time were dictated by the compromise of the BHC Act, which had left many loopholes intentionally and also contained enough ambiguity to allow others to form. For example, in 1957 the OCC consented to a minor erosion by allowing banks to profit from providing their customers brokerage transactions, an activity previously barred from any profit, although the regulator still insisted that such services only be offered to existing customers rather than marketed to the public (OCC 1957). Such incremental decisions represent the most common form of erosion.

Courts sometimes arrested these small steps toward expanded powers for commercial banks, though. In *Baker, Watts & Co. v. Saxon* (D.D.C. 1966) a federal district judge invalidated the Comptroller’s decision to allow a bank to underwrite and trade state bonds not backed by the taxing power. Glass-Steagall itself contained an exception for “obligations of the United States, or general obligations of any State or of any political subdivision thereof,” but the court interpreted “general obligations” narrowly, to include only those debts directly backed by the state’s power of taxation, thus preventing this intentional loophole from growing.

In 1965, Congress took aim at many of the loopholes left by the 1956 Act. House Banking Committee Chairman Wright Patman (D – TX) first introduced a narrow bill repealing one particular exemption for testamentary trusts, which had allowed the blossoming of a miniature empire of unregulated firms in Florida under the auspices of the estate of Alfred DuPont. But Rep. Charles E. Bennett (D – FL), encouraged by the Fed, offered a comprehensive
amendment to level the playing field, which would have repealed all of the most important exemptions, including for one-bank BHCs. This broader maintenance bill passed the House in spite of opposition by Patman, who argued that every exemption deserved individual consideration before being repealed (CQ 1965, 856-7). Although the Fed was adamantly in favor of ending the one-bank exemption, doing so was strongly opposed by the IBAA, ABA, and OCC (who said “imaginary possibilities of abuse” were no reason for action and worried of Fed jurisdictional imperialism). To secure passage, the Senate Banking Committee dropped the one-bank exemption before sending the bill to the floor (CQ 1966, 762-766). The House assented to the Senate amendments by voice vote, and the bill was signed into law by a fairly uninterested President Johnson on July 1, 1966, having closed many important loopholes but leaving the most important one untouched. Most legislators were apparently still committed to a separation of commercial and investment banking, willing to support loophole-closing upkeep but hoping to avoid angering any constituency while doing so.

Their political calculations in 1965 seem to have been fairly short-sighted, as rapidly growing numbers of BHCs took advantage of the one-bank loophole to escape Fed supervision. In 1965 there were 550 one-bank BHCs, by year-end 1968 691, and by February 1969 810, controlling 40% of all commercial bank deposits (CQ 1969, 942). President Nixon urged Congress to stop this “disturbing trend in the past year toward erosion of the traditional separation of powers between [banks and commerce].” Chairman Patman now abandoned his defense of the one-bank loophole, and indeed derided the Administration for being too lax. After

---

11 CQ 1969, 941-2 provides a list of other loopholes being used. The desire to escape regulation leads to some fairly ingenious means of evasive action, some of which seem obvious in retrospect but many of which simply seem impossible to anticipate. An obvious one was the minimum of 25% ownership of a securities firm that triggered regulation. BHCs simply found ways to exert controlling interests with 24% of shares. This loophole was closed by the 1970 amendments by simply allowing the Fed to determine when a controlling interest was present, showing the need for flexibility to achieve policy results robust to private attempts at evasion (CQ 1970, 874).
struggling over who should have regulatory jurisdiction and the power to define the scope of permitted activities (and therefore who would determine the responses to future erosion), Patman’s Fed-centric bill ultimately won out in the House. The Senate passed an amended version of the bill, but not before adding several exemptions for very small BHCs and those in which banking was only a small part of the company’s overall operations, as well as clarifying that “traditional banking practices” were not meant to be targeted (CQ 1970, 880). The ABA intensely lobbied House conferees to accept these bank-friendly Senate Amendments and mostly succeeded. Once again, loopholes that served well-positioned interests were included as long as they seemed not to threaten the overall efficacy of the law.

Importantly, there were some unresolved disputes between the House and Senate on how the act’s language would operate. Although Congress had considered including a “laundry list” of prohibited activities, in the end they failed to nail one down and instead legislated the rather vague standard that any allowed nonbank activities must result in public benefits (Jessee and Seelig 1977, 30-37). In practice this left the Fed to exercise its own discretion to determine what “public benefits” would mean, as courts would be unable to discern a clear congressional intent capable of trumping reasonable bureaucratic interpretations. Congress may be capable of providing the surest maintenance, but doing so may require a degree of political consensus that seldom exists. Remembering the tradeoff between maintenance and political costs, often achieving this consensus may not be worth the effort, and so the responsibility of maintenance is likely to be shunted onto regulators—with the added political bonus of plausible blame-avoidance if anything should go so wrong as to attract critical public attention (see Fiorina 1989).

The next major developments for Glass-Steagall and the BHC Act came in the federal courts. Courts’ centrality was guaranteed by the unanimous holding in Association of Data
Processing Service Organizations v. Camp (1970), which established the precedent that aggrieved competitors were entitled to standing to challenge regulatory decisions in spite of the absence of any explicit statutory provision providing it (156-7).12 Thenceforth, the chief opponents of new powers for banks—their potential competitors—would be able to bring legal challenges, an opportunity they would not waste. Securities groups quickly put their standing to use in Investment Company Institute v. Camp (1971). The ICI, a securities firm trade group, argued that the Comptroller’s decision permitting banks to offer customers “investment funds” violated Glass-Steagall’s ban on commercial banks conducting securities activities. The District Court ruled against the Comptroller, but the Court of Appeals reversed, deferring to the Comptroller. Writing for a Supreme Court majority, Justice Stewart acknowledged the need for deference, but found that, whatever the Comptroller might say, the investment funds in question were so nearly indistinguishable from mutual funds as to represent violations of the Glass-Steagall Act (625-7). Deference was inappropriate when the Comptroller had offered little rationale for its holding (with appellate counsel’s “post hoc rationalizations” deemed too little, too late) and when Glass-Steagall’s legislative history clearly indicated that mutual fund activity was meant to be prohibited (628, 631-3). In light of these facts, Justice Stewart concluded that the Comptroller’s narrow reading of the term “securities” would vitiate the enacting Congress’s will and was thus impermissible (634-6). In short, the Court found that the OCC’s erosive interpretation was unjustified on both textualist and purposivist grounds and replaced it with a maintaining interpretation. Justice Blackmun dissented, complaining that the majority’s opinion was “based more on what is…desirable national banking policy than on what is a necessary judicial construction of the Glass-Steagall Act…Policy considerations are for the Congress and

---

12 The case centered on powers granted to national banks by the OCC under the National Bank Act rather than directly implicating Glass-Steagall or the BHC Act, but the principle of its holding was easily applied to these closely related laws.
Blackmun thus found textualism inadequate to settle the ambiguities, and preferred to defer to the Comptroller’s expertise as to how this particular decision would serve the function of the law. Once again we see the Justices arguing about the proper means of achieving consistent law, with the majority’s concern for a gap between legislation and bureaucratic actions this time leading it to undertake active maintenance.

D. A Change in Commitments: 1970s to 1980s

The next phase of Glass-Steagall’s development features a weakening of the commitments to the law’s goals. At the same time, there was little legislative push for deregulation or reform. Congress was too internally torn to move toward fundamental shifts; bureaucrats sought to interpret existing laws to expand their own supervisory powers while blurring the boundary between commercial and investment banking; and judges, confronted by a rising tide of security industry complaints, found that they were the ones charged with figuring out what sorts of maintenance were necessary.

Congress had taken care in 1970 to give administrative responsibility for Glass-Steagall and the BHC Acts to the Fed rather than dividing it among the three banking agencies, presumably in hopes of keeping implementation strict. It came as a surprise to legislators, then, that the Fed from 1970-1974 turned out to be fairly lenient in allowing non-bank activities, leading the BHC became far and away the dominant corporate form for commercial banking during the period. In 1972 the Board amended the regulations governing which actions were “closely related to banking,” enlarging the permissible scope of activities for BHCs (Bd. Of Governors v. ICI, 1981, 48). Only in 1974, with an economic downturn, did the Fed return to its conservative temperament, adopting a “go slow” policy and multiplying its denials of applications for new activities (Jessee and Seelig 1977, 34, 39). Meanwhile, the Comptroller
continued to facilitate slow but steady erosion, for example in 1974 interpreting Glass-Steagall to let banks advertise computer-assisted stock purchasing services, although they could still only be offered to those with checking accounts at the bank (OCC 1974).

By this time, four decades after the passage of Glass-Steagall, banking regulators no longer prioritized maintenance of a strict wall between commercial and investment banking, and their actions would become increasingly erosive in the proceeding decades. Understanding their motivations for this shift is crucial to understanding the remainder of Glass-Steagall’s history. In the deregulatory era of the late 1970s and 1980s, it is easy to imagine that regulators were simply caught up in a larger intellectual movement (see Derthick and Quirk 1985)—whether one believes this was a great awakening or the historical apogee of regulatory capture. But officials’ thinking at the OCC, FDIC, and Fed was at least somewhat more complex than this explanation would suggest.

Over the course of the 1970s, the nation’s complex system of commercial banking regulations was in danger of being marginalized as Wall Street securities firms offered attractive alternatives to commercial banks’ services. Corporate bank loans were increasingly displaced by commercial paper (short-term corporate bonds), which provided a cheaper means for corporations to raise funds. Bank deposits, subject to strict interest rate ceilings until early-1980s legislation (discussed shortly), had to compete with the new, more flexible money market mutual fund industry. This phenomenon of commercial banks being cut out of the loop, called “disintermediation,” meant that the well-developed regulatory system for commercial banking covered a shrinking proportion of the nation’s assets, a trend which pleased the securities industry but worried banking regulators charged with overseeing the stability of the nation’s financial system (Chernow 1997).
As a result, the Fed hoped to expand the menu of services that BHCs could offer, and the OCC hoped to give national banks the broadest powers that their statutes could support. By allowing them access to profitable activities, they hoped to strengthen the overall safety and soundness of their regulated institutions, increasing their ability to withstand difficult economic times. This goal was reflected in regulators’ joint emphasis on capital adequacy requirements, and was also made explicit in 1987 when the Fed announced it would require affiliated banks to channel revenue toward any troubled banking affiliate, thus making the expanded scope of banks’ activity a source of systemic strength (Fed. 1987b). Regulators were thus largely motivated by a concern that animates banking policy debates today, as well: the desire to avoid taxpayer-funded bailouts of large financial institutions. Embodying this position, Comptroller of the Currency C.T. Conover delivered a speech in 1984 entitled, “Strength = New Powers + Firm Supervision” (OCC 1984).

That said, when it comes to specific Glass-Steagall prohibitions on mixing commercial banking with other functions, the bureaucratic interpretations of the 1980s were decidedly erosive. Conover declared that the OCC’s position was that “if a bank can make a strong case that a proposed activity is legal, our inclination is to approve it” (1983). The OCC abandoned limitations on banks’ brokerage services (1982) and opened the door to bank subsidiaries marketing insurance (1986). Acting in parallel, the Fed allowed bank subsidiaries to market and underwrite all types of government securities (1984b), provide full brokerage services (1986), and to trade or underwrite revenue-bonds and mortgage-backed securities (1987d). Affiliates were given full power to deal in commercial paper (1987c). Bureaucrats, then, had become complicit in the process of eroding Glass-Steagall’s wall of separation between commercial and
investment banking functions. Where once they had given their fingers to plug holes in the dike, they now removed them in ways that allowed commercial banks to expand their powers.

Should we understand this shift simply as regulators going over to the other side—being so thoroughly captured by industry that they conceived of their own purposes as entirely overlapping with their regulated firms’ interests? There is undoubtedly some truth in this view, but it fails to sufficiently appreciate the dilemma that bureaucrats felt they faced. If we view their highest priority as realizing the Glass-Steagall Act’s separation, thereby faithfully honoring the vision of the statute’s framers of the 1930s, then it is fair to say that the regulators were neglecting their duties. On the other hand, if we think of these regulators as owing a deeper fidelity to protecting a functional and well-regulated banking sector through use of the larger edifice of banking laws, their actions can make sense without any cynicism. Regulators thought they could maintain the vitality of the law’s deeper spirit by allowing Glass-Steagall’s requirements to bend, whereas they believed that a literalistic fidelity to the law would lead to a disastrous break in the whole system. Acting as faithful agents to a long-dead legislative principal was not their first priority; maintenance on one front thus led to erosion on another.

During these years, certain types of erosion—those directly provoking important constituents—still aroused some members’ instincts to maintain the act but never produced legislation. The House passed a bill in 1980 to restrict BHCs from acting as insurance agents, but the Senate never responded. The large banking bill passed that year, the Depository Institutions Deregulation and Monetary Control Act of 1980, decontrolled interest rates for deposits, ratifying erosion of long-standing policies that the banking regulators had sought but the courts had denied in ABA v. Connell (D.C. Cir. 1979). Similarly, another broad reform law, the Garn-St. Germain Depository Institutions Act of 1982, contained few provisions relevant to
Glass-Steagall in spite of Senate Banking Committee Chairman Jake Garn’s (R-UT) desire to address the issue (Filipiak 2011, chap. 4).

Signs soon emerged that Congress was considering abandoning the project of maintenance and turning toward large-scale “reform” of the act, which would mean abandoning some of the central provisions. In 1983 Ronald Reagan proposed Glass-Steagall reform, including allowing banks to conduct real estate operations, but it was not an administration priority and came to naught. The House considered legislation in 1984 to close remaining loopholes in the 1956 legislation as the Senate worked at cross-purposes, passing a bill to allow BHCs to form subsidiaries dealing in mortgage-backed securities and municipal bonds. No legislation resulted (CQ 1999, 5-26-5-27).

Even as regulators’ altered their focus and Congress failed to change the law to keep up with this change, the law itself continued to act as a significant constraining force. The bureaucrats themselves were not immune to textualist misgivings about offering expanded powers without changing the laws. Several decisions made by the Federal Reserve split the Board of Governors, 5-2, with rare dissents from Chairman Paul Volcker, whose objections were based less on policy disagreement than on a lack of legal authority to approve the changes under existing law. Volcker wrote: “As the Board as a whole has repeatedly urged, the plain and desirable remedy to this legal and substantive morass is a fresh congressional mandate” (Fed. 1987c, at 506). Volcker and other Fed officials also testified in numerous congressional committees throughout the early 1980s, practically begging legislators to address “the significant erosion, at the margin, of the commerce-banking barrier” and other developing legal rifts (e.g.

---

13 Importantly, many policymakers viewed maintaining the separation between banking and commerce as far more important than the separation between commercial banks and investment banks. Representative Jim Leach (R – IA) cites maintenance of this separation between banking and commerce as one of the primary motivators for reformers in the 1990s (Leach 2008).
Fed. 1983, 1984a, 1987a, 1987e). Central to these pleas was the argument that the Fed was incapable of doing the right kinds of legislative maintenance without new statutory language. As it was, the Fed could provide legally creative (and sometimes dubious) patchwork fixes subject to judicial second-guessing, providing at best uncertain resolution.\footnote{Such institutional patchwork recalls the “state building as patchwork” described in an earlier era by Skowronek (1982). Just as building America’s administrative state as a whole proceeded through an incremental reconstruction of existing institutions, so too is the same messy, incremental, and layered process played out in the context of reconstituting statutory institutions capable of pursuing particular policy goals.}

With regulators having largely abandoned maintenance of the original goal of total separation, and Congress having become somewhat ambivalent about the importance of maintaining the wall, the task of maintenance in the 1980s fell mostly to the courts. If we suppose, as is sensible, that they had no special policy preference in favor of preserving Glass-Steagall, then their commitment to maintenance derived from their generic desire to preserve some respect for the laws on the books, thereby giving force to Congress’s choices and creating a predictable and coherent environment for regulated banks. This sometimes led them to take a hard line on upholding statutory language, but not always, as courts balanced their preferences for textualist maintenance against expert regulators’ calls for flexibility.

In \textit{Board of Governors v. ICI} (1981), the Supreme Court enabled erosion by deferring to the Fed’s determination that banks could affiliate with closed-end investment companies, distinguishing the case from \textit{ICI v. Camp} because the regulator offered an extensive, convincing opinion and the fact that affiliates, rather than banks themselves, would be involved (68). Continuing the trend toward deference, the D.C. Circuit ruled in favor of the Fed’s decision to allow a state member-bank to deal commercial paper in \textit{A.G. Becker Inc. v. Board of Governors} (1982). The Fed now interpreted the Glass-Steagall Act’s prohibition on banks marketing “securities” to exclude commercial paper, arguing that these instruments functioned enough like
loans to fit into the core functions of banking (139). Showing sympathy for a regulator trying to make the best of this situation, the majority in *Becker* emphasized the deference due to the Fed’s opinion because of the agency’s expertise, the lack of a precise definition in the statute (“we cannot assume that Congress intended the term to comprise a set of rigid and unchanging categories”), and the thoroughness and validity of the board’s reasoning (140-1). The court thus endorsed the Fed’s functional analysis, agreeing that commercial paper was more like normal bank lending than securities speculation (149). A vigorous textualist dissent argued that the majority’s holding was inconsistent with *ICI v. Camp* and contrary to Glass-Steagall’s simple language (152).

The overall trend toward deference to agencies’ functional interpretations was reinforced by *Chevron v. NRDC* (1984), the landmark case involving EPA regulation of factory pollution. Justice Stevens, writing for a unanimous Court, set out a two-pronged test for agency deference fated to become canonical: first, if the intent of Congress is clear in the statute’s language, this is always definitive for courts; second, if ambiguity exists, an agency’s reasonable and permissible constructions of the statute deserve judicial deference (842-3). *Chevron* belonged to a long-term trend toward deference, but its independent causal influence would grow as courts worked to sort out the banking laws in the years to come.

When regulators became too aggressive and appeared to simply ignore the statute’s plain meaning, however, courts were still willing to take a stand. A string of cases in the years after *Chevron* probed the limits of deference. Two different strands of cases are particularly important. The first had to do with the Federal Reserve’s ability to define what constituted a bank for the purposes of the BHC Act. The second returns to the issue of commercial paper debated in *A.G. Becker*, discussed above.
In *First Bancorp. v. Board of Governors* (Feb. 1984), the Tenth Circuit showed that deference was not absolute, in this case ruling that erosion could not be halted without some statutory sanction. “Banks” under the BHC Act were defined as institutions offering both demand deposits and commercial loans, and the early 1980s saw a rise in so-called nonbank banks, institutions that offered only one of these services and thus circumvented regulation. Pernicious loophole or not, though, the court found that the Fed’s sudden decision to treat as a bank any entity that offered both commercial loans and negotiable orders of withdrawal (NOW) was clearly inconsistent with the unambiguous statutory definition of what was to be considered a bank (436). To reach its preferred outcome, the Fed had to treat NOWs as creating a “legal right of withdraw on demand” and thus ignore their specific legal terms that made such withdrawals “negotiable.” Such a functional flexibility in interpreting the law might have been sensible, but the court insisted it illegitimately created an unpredictable legal environment for regulated actors doing their best to comply.

The Fed did not relent, promulgating a new official definition of a bank affecting both NOW accounts and commercial loans, this time offering more extensive justification. It quickly found itself back in front of a Tenth Circuit panel in *Dimension Financial v. Board of Governors* (Sep. 1984), where it argued that non-bank banks enjoyed an unfair competitive advantage by escaping Fed regulation (Sup. Ct., *Bd. Of Governors v. Dimension Financial*, 1986, 367). In reviewing the Board’s change in the definition of commercial loans, the court declared that there had been no changes in common usage, congressional directive, or regulatory usage regarding the meaning of “commercial loan” since 1970 other than the Fed’s new regulation. It then chastised the Fed for making changes “with no reference to the actual meaning,” but rather “purely…to stop changes in the business of providing financial services—thus to bring a large
number of additional enterprises under its jurisdiction.” The judges further scolded this “device
to freeze the changes in the business of financial services until the Board could persuade
Congress to act” (10th Cir. 1984, 1405). The court reasoned that the institutions the Fed now
sought to target had perfectly legally exploited “permitted exceptions”—intentional loopholes—and
could not now suddenly be treated as if they had impermissibly evaded the act, and was
emphatic that if anyone was going to fix this “problem, if it is that,” it would have to be
legislators, who had proven themselves perfectly capable of performing statutory maintenance on
the BHC Act in the past (1407-8).

A unanimous Supreme Court affirmed in Board of Governors v. Dimension Financial
Corporation (January 1986), following the lower court’s logic closely. Chief Justice Burger
found that Chevron’s first prong applied, since the statutory language creating the loophole was
unambiguous and therefore did not leave the Federal Reserve latitude to close it at will (368).
The Board, in its maintaining manifestation, justified its position by arguing that “a statute must
be read with a view to the ‘policy of the legislation as a whole’ and cannot be read to negate the
plain purpose of the legislation” (373). The Court, however, declined to entertain this
purposivist logic, insisting that such imperfections in the statute were Congress’s to fix (374).

At stake in the Dimension cases was whether Congress could be cut out of policymaking
in an area dominated by experts. If the courts had assented to the Fed’s policy changes, they
would have effectively given the banking agencies unlimited scope in deciding how and when to
perform policy maintenance, thereby depriving Congress of ultimate authority to shape the
functioning of particular statutes. As with Justice Blackmun’s 1971 ICI v. Camp dissent, the
Court showed remarkable faith in Congress’s ability to act over long expanses of time to
maintain statutes. Subsequent cases would expose this faith as limited, though.
The second strand of cases, which began in 1982 with the D.C. Circuit’s above-described A.G. Becker decision, temporally overlaps with the first. In its June 1984 case known as Becker\textsuperscript{15}, the Supreme Court overturned the D.C. Circuit. Justice Blackmun, writing for the majority, insisted that commercial paper “falls within the plain language of the act,” and therefore should be treated as a security for the law’s purposes (140-1). The Fed’s preference for functionally justified erosion in this context was deemed illegitimate, as it required disregard for statutory language (141-2). The Court cited its ICI v. Camp decision in support of various parts of its logic (144-145), which is especially ironic given that the majority in the older case (which Justice Blackmun had not joined) had embraced a functional reading of Glass-Steagall in contrast to the Comptroller’s narrow, erosion-enabling reading.

Becker shares with Board of Governors v. Dimension Financial (to be decided by the Supreme Court nineteen months later) an insistence that the Fed is not allowed to treat itself as ultimate arbiter of how to “make the statute as a whole work.” Instead, Congress must guide the course of erosion and maintenance, with its inaction representing an “adherence” to previous statutory language (Becker, 153). This logic would suggest that ICI v. Camp was decided wrongly, as Justice Blackmun originally argued. Foreshadowing things to come, Justice O’Connor, joined by Justices Brennan and Stevens, dissented in Becker. Because there was no explicit definition of securities provided, the dissenters would have deferred to the Board’s determinations on grounds of expertise, in line with Chevron (handed down three days earlier).

In the wake of Becker, courts attempted to figure out what to do with the Comptroller’s similar decision to allow banks to establish individual retirement accounts (IRAs). ICI v.

\textsuperscript{15} This case’s formal name is SIA v. Board of Governors, but is often referred to as Becker in other court cases and in the literature. Alternatively, this line of cases is sometimes called Banker’s Trust, after the name of the bank seeking to market commercial paper (whereas Becker and the SIA represented the securities firms whose competitive interests were damaged by the Fed’s permissive decision).
Conover (N.D. Cal. Aug. 1984) aligned with Becker, insisting that the Comptroller was charged with administering unchanging statutory language and not with creating its own discretionary law (854). This was to be the high-water mark for the textualist position embodied in Becker, though. In a parallel action, also called ICI v. Conover (D.D.C. Nov. 1984), the ruling instead followed Chevron in finding the Comptroller’s IRA decision worthy of deference, and the same result obtained in another nearly identical action, ICI v. Clarke (D.Conn. Jan. 1986). The Second Circuit affirmed the deferential result in a quick per curiam in May 1986, and a unanimous D.C. Circuit followed suit later in the month. Noting the extensive justification provided by the Comptroller in making his ruling, the D.C. Circuit distinguished Camp, at which point it could point to ambiguous statutory language and legislative history and rely on Chevron deference (though nothing had really changed since 1971) (933-935). Rather than seeing statutory deficiencies as being Congress’s to correct, the court wrote that “Congress has charged the Comptroller with making” complex economic judgments “in the first instance,” and that it was none of the courts’ business to interfere with that delegation (938). The Ninth Circuit overturned the outlying Northern District of California decision in ICI v. Clarke (June 1986), declaring that “whether [the Comptroller] performs his task wisely is not for us to decide” (222). The Supreme Court denied certiorari in November 1986.

Finally, capping this strand of cases, the D.C. Circuit took up another case called Securities Industry Association (SIA) v. Board of Governors (Dec. 1986). A unanimous three judge panel challenged Becker head-on, upholding a permissive commercial paper ruling made by the Fed on remand from Becker that circumvented the Supreme Court’s specific complaints by referencing a different section of Glass-Steagall (1055). Noting that the Board had now “comprehensively addressed the language, history and purposes of the Act,” the court found the
action met *Chevron*’s deferential standard: the interpretation could be reconciled with the statute’s text and was reasonable (1067). Once again, this time perhaps more surprisingly, the Supreme Court denied cert in June 1987, choosing to remain silent as its own precedent was distinguished into oblivion. This was a quiet but effective way for the Court to allow *Chevron* deference to become the controlling rule in banking regulation without suffering the embarrassment of directly reversing one of its own decisions.

It seems likely that as the 1980s wore on and congressional attention failed to be converted into congressional action, the textualist insistence in the legislature’s ability to perform maintenance no longer seemed to be based on plausible assumptions. Holding the line until Congress acts, even as doing so frustrates bureaucrats’ efforts to make the overall system of regulation function sensibly, became a difficult proposition to defend. In the persistent absence of renewed congressional action, judges’ limited capacity and resources made acquiescence to the banking regulators the court’s best hope for achieving a consistent, sensible legal environment for banks. In such a fast-changing context, achieving the coherence aspect of maintenance was in tension with the purposive aspect.

Though judges might hope to at least secure coherence by letting the executive branch have its way, the resolution that bureaucrats could provide was in many ways problematic, being both uncertain and jurisdictionally messy. With the recent commercial paper holdings allowing banks to act so much like securities firms, the SEC asserted its right to regulate them and convinced a D.C. District Court judge of its power to do so in *ABA v. SEC* (Nov. 1986). A unanimous D.C. Circuit panel reversed, stating that although “the original administrative construction of the Glass-Steagall Act has been dismantled piecemeal over the last fifty years” this did not justify the SEC’s purposivist argument for jurisdiction, which had no basis in the text.
of laws which made jurisdiction a function of “the species of financial institution.” The SEC may have been espousing sensible policy, but it was inarguably unsupported by the Glass-Steagall or Securities Exchange Acts (741-3). Judge Wald sympathized with the SEC’s argument from functional need, but nevertheless asserted that “it is not the job of courts to play the game of ‘what if’ in predicting hypothetical congressional reactions to hypothetical circumstances,” especially when the statutory language itself was so clear (749). Showing worries for courts’ capacities, Wald wrote that doing so “would embroil courts in constant controversies as to whether the context that had changed was really relevant” (754).

Erosion continued with regulator encouragement now, for example with the FDIC’s decision to grant the state banks it regulated the ability to have securities affiliates. The D.C. Circuit upheld this decision in *ICI v. FDIC* (1987), finding that Glass-Steagall’s language was clearly meant to give regulators the power to make such a choice, and therefore *Chevron*’s first prong analysis was all that was necessary to decide the question (1546).

**E. Struggling Over a New Regime: Late 1980s to 1999**

With so many holes in the floodwall and so much water over the dam, one might think that Congress would feel compelled to assert its primacy by at least determining how erosion would continue. In the decade from 1988 to 1998, broad legislative coalitions made many attempts to do so, but consistently came up shy of their goals. In large part, this was because bureaucratically engineered erosion during the 1980s had lessened commercial banks’ desire for broad legislation, as they could afford to oppose any changes in the statutes that failed to match the steadily improving treatment that changed interpretations had already delivered.
Congress seemed poised to provide maintenance in 1987, when it brought non-bank banks into the BHC regulatory framework in the (mostly unrelated) Competitive Equality Banking Act of 1987. It simultaneously signaled its intent to take larger action against regulatory erosion by imposing a one-year moratorium on regulatory approval of BHC applications to engage in real estate, insurance, or securities underwriting activities, giving legislators time to draft a more ambitious proposal. That came in 1988 in the form of the bipartisan Proxmire-Garn bill, negotiated by Senate Banking Committee Chairman William Proxmire (D – WI) and Ranking Member Garn, which would allow banks to market securities subject to Fed supervision and “firewalls” prohibiting risky arrangements, as well as conferring a limited ability to market insurance. With ABA and insurance industry support, the bill sailed through the Senate, 94-2 (CQ 1987, 632; CQ 1988, 231-2). In the House, however, Glass-Steagall reform became entangled with key legislators’ other priorities. Banking Committee Chairman Fernand St. Germain (D – RI) insisted on including strict consumer protections and an expansion of the Community Reinvestment Act (CRA); House Energy and Commerce Committee Chairman John Dingell (D – MI), a champion of the securities industry, wanted to sharply limit bank affiliates’ permitted activities (CQ 1988, 236, 241). No compromise was reached, and the congressional session ended without House action.

At this point, new Fed Chairman Alan Greenspan wrote and repeatedly testified to urge action, with a thinly veiled message that the Fed would act on its own if Congress would not. The outgoing Proxmire (headed to retirement deprived of a career-capping achievement) supported Greenspan’s message, but St. Germain was emphatic that he would not be dictated to by the Fed, saying he would feel free to overturn Fed actions (CQ 1988, 241)—a hollow threat, since what the Fed was demanding was congressional action.
With the fallout from the savings and loan crisis dominating the Banking Committees’ attention in the years to come, St. Germain never made good on his promise. The giant Financial Institutions Reform, Recovery, and Enforcement Act of 1989 failed to address banking powers except to empower banks to acquire distressed thrifts (CQ 1989, 122). Erosion by the Fed and OCC continued apace. In November the Fed allowed banks to derive as much as 10% (up from 5%) of their revenue from their subsidiaries’ securities activities (1989). This had an immediate impact: total bank-ineligible securities underwritten by these subsidiaries nearly doubled from $36.2 billion in the third quarter of 1989 to $68.7 billion in the fourth quarter, and by 1991 these firms accounted for 10% of the market for certain types of financial products, including asset-backed securities (GAO 1990, 1991). By 1995, they would represent nearly 15% of all broker-dealer assets, meaning some of the most important firms in the securities business were being regulated as if they were only commercial banks—a problem the GAO diagnosed as endemic to congressional inaction, and not eradicable by the regulators themselves (GAO 1995).

In 1991, at the urging of President Bush, Congress took up Glass-Steagall reform again alongside an overhaul of the nation’s acutely distressed deposit insurance system. Even the securities industry became cautious supporters of reform, wanting reciprocal access to banking activities instead of defending an increasingly illusory strict separation. Reform would be useful to them only if it enabled banks to affiliate with non-bank commercial enterprises, however, since most controlled non-financial subsidiaries (CQ 1991, 76). Although the House Banking Committee passed a broad bill that resisted most attacks, including Jim Leach’s (R – IA) amendment to prevent affiliations between banking and commerce, it had to subsequently

---

16 This reversal of the securities industry’s position nicely illustrates the opportunistic orientation of interest groups toward erosion or maintenance. This dynamic closely parallels what Hammond and Knott (1988) describe (with a metaphor contrary to this chapter’s) as the “deregulatory snowball.” They argue that when the status quo of regulation is sufficiently disrupted, those who found themselves as beneficiaries of regulation in the past may now find themselves receiving the short end of the stick and call for the abandonment of regulation altogether.
face the demands of four other committees, where opposing interests bogged it down (CQ 1991, 84). In the end, Glass-Steagall reform was thrown overboard to secure a narrow consensus on deposit insurance reform (CQ 1991, 89, 91).

The Republican takeover of 1994 gave Glass-Steagall reformers new hope, but it was to prove unfounded. Regulatory erosion had lessened the urgency felt by many commercial bankers, who were already doing much of what they wanted without having to change the underlying statutes. Some members of Congress, including new House Banking Committee Chairman Leach, were still determined to “bridge the gap between law and reality,” but once again other priorities intervened. Although Leach’s committee easily passed his reform bill, resistance from the Commerce Committee and disinterest in the Senate doomed the effort once again. In an attempt to restore momentum, reform was married to regulatory relief, along with a permanent moratorium on OCC expansion of bank powers. This last provision would prevent future erosion directed by the banks’ favored regulator, but it soured commercial bankers on the entire package and the nascent compromise soon unraveled (CQ 1995, 2-80).

Commercial bankers could afford to reject legislative action in part because they (rightly) expected the Supreme Court to ratify OCC decisions confirming some bank insurance powers (CQ 1995, 2-84). They were not disappointed: unanimous decisions in 
Nationsbank v. Variable Annuity Life Insurance Co. (1995) and Barnett Bank v. Nelson (1996) paved the way for BHC affiliates to sell insurance by deferring to the OCC. The ruling allowed banks to greatly expand their insurance operations and thus further weakened their demands for legislation.

Back in the legislative arena, in 1996 Chairman Leach continued to press for Glass-Steagall reform, but was ultimately compelled to table it in favor of other priorities again.
As other bills gained momentum, he tried to perform a small act of Glass-Steagall maintenance by attaching a provision subjecting banks that marketed insurance to state licensing requirements, but the newly empowered banks saw the provision as undercutting their gains from regulators and courts and therefore opposed it, and Leach’s efforts stalled again (CQ 1996, 2-44-2-54; McConnell 1996).17

After these legislative failures, the Fed acted on its own volition to allow BHCs to expand securities operations in affiliates, changing the cap on affiliate security revenue from 10% to 25%. Similarly, the OCC began allowing bank subsidiaries to perform activities previously restricted to affiliates (CQ 1996, 2-54). These changes meant that the legislative status quo now advantaged commercial banks at the expense of the securities and insurance industries. Glass-Steagall reform became a priority for these industries, who found that banks were restricted from entering their industries only where the original statute was crystal-clear, while they were still barred from engaging in banking activities (CQ 1997, 2-74).

In 1997 legislative efforts resumed yet again. Chairman Leach’s goal was to repair the now woefully fragmented, bank-favoring regulatory edifice, providing maintenance of the congressional goal of having a fair, coherent, and capable regulation even as the erosion of the wall between commercial and investment banking continued. Reform would provide an “equalitarian” framework, while also still maintaining the separation between banking and commerce (Leach 2008). Early in the year, a merger between Bankers Trust and the brokerage Alex. Brown had been approved, representing the biggest and most blatant breach of Glass-

\[17\] A regulatory relief bill passed that year contained only a minor Glass-Steagall provision. Well-capitalized BHCs seeking to expand into non-banking activities approved by the Federal Reserve had previously needed to apply and wait up to 90 days for Fed Board approval, but could now proceed just by notifying the Fed within 10 days of commencing the new activity (CQ 1996, 2-49). Congress was fully cognizant of the dynamic at work and here streamlined erosion even as it failed to provide more sweeping reform. Apparently, the interest group dynamics that made it too politically costly to achieve wholesale change did not prevent some harmonization of law and practice.
Steagall yet. In May former Fed Chairman Volcker told Congress that legislation was overdue, with the resolution provided by courts and regulators being uncertain and inadequate. Members of Congress other than Leach began to feel jealous of their own institutional prerogative, as well. Conference Chairman John Boehner (R – OH) declared, “Congress wants to preserve our right to write laws, which frankly has been overshadowed by the OCC and other regulators” (CQ 1997, 2-75). The legislative wrangling during 1997 bogged down yet again, however, as there were heated intraparty disagreements about whether to allow a mix of banking and commerce (which Leach vehemently opposed). A compromise brokered by the Republican leadership in March 1998 stalled yet again when the ABA deemed the compromise unpalatable. Leach sought to lure commercial banks back to the table by including a provision to allow limited securities powers to bank subsidiaries (rather than BHC affiliates), and hoped to gain Democratic support by tying in an extension of the CRA (CQ 1998, 5-5-5-8).

At this point the political landscape was altered by the announcement on April 6, 1998, of the merger of Citicorp and Travelers Group, at that point the biggest stock swap in history. Both companies’ shares surged as the intended merger was greeted with enthusiasm, in spite of then-existing laws requiring (in no uncertain terms) that the merged company divest itself of insurance and underwriting activities—which accounted for about half of Traveler’s Group’s revenue. The Fed approved the merger, giving a five year window for divestment—though most expected that congressional action would finally arrive before then (CQ 1999, 5-23).

With more pressure on Congress than ever before, Leach’s reform bill was reintroduced in the House in May 1998. Leach successfully proposed an amendment stripping the provisions that would allow the mixing of banking and commerce, 229-193 (CQ 1998, 5-9-5-11), and the final bill squeaked by 214-213, backed by about 2/3 of Republicans and 1/3 of Democrats after
many favors were called in. Senate Banking Committee Chairman D’Amato reluctantly agreed to try to get a bill through his chamber, but this attempt ran into Senator Phil Gramm’s (R – TX) intransigent opposition to the CRA (CQ 1998, 5-10-5-12).\textsuperscript{18} Meanwhile, Treasury Secretary Robert Rubin guaranteed a veto if the bill invested supervisory powers wholly in the Fed rather than the Treasury Department—indicating what many called a “turf war” (CQ 1998, 5-14). Despite the breadth of the interests favoring compromise, 1998 ended without Senate action.

In 1999 the new Congress picked up where the old one left off, with Gramm now Chairman of the Senate Banking Committee after D’Amato’s November defeat. Finally, this was to be the year. Gramm’s Senate version would allow non-banking subsidiaries for banks with assets less than $1 billion and was unsurprisingly uncompromising on the CRA. With its controversial provisions, it passed the Senate floor on a party-line vote, 55-44. A broader coalition was forged in the House by grafting elements of a Commerce Committee effort onto Leach’s Banking Committee bill, and in spite of fierce floor battles over peripheral issues, the House bill won the administration’s backing and strong bipartisan support, passing 343-86 on July 1 (CQ 1999, 5-5-5-15). Conference committee negotiations featured intense conflicts, especially over the CRA, where Gramm doggedly held out until he won concessions. The Fed and Treasury eventually announced a truce in their jurisdictional disputes, agreeing on a structure that gave the Treasury much of what it wanted on bank subsidiary powers (CQ 1999, 5-25). After more than a month of wrangling, a compromise emerged that won passage in both houses on November 4, 363-57 and 90-8, respectively. A week later President Clinton signed into law the Financial Services Modernization Act (Pub. L. No. 106-102), better known as Gramm-Leach-Bliley (GLB) (CQ 1999, 5-26, 5-31).

\textsuperscript{18} It should be noted that Gramm’s intense opposition to the CRA appeared to be ideological rather than interest-group-driven. At this point in time, the ABA’s representative said that banks did not feel particularly strongly on the issue (CQ 1999, 5-10).
Most sources regard GLB simply as a repeal of Glass-Steagall, and understandably so. A new corporate form was born—the Financial Services Holding Corporation—that explicitly allowed commercial banks, investment banks, and insurance companies to reside under the same corporate umbrella. Representative Leach’s office held a wake for the old law, which featured a cake frosted with “Glass-Steagall, R.I.P., 1933-1999” (“Washington People…,” 1999). As we have seen, while eulogies may well have been in order, “repeal” of the law’s substance had mostly been accomplished by a long process of erosion. Representative Leach thought of the act as “procompetitive, but not anti-regulatory” reform (Leach 2008).

Postscript: Glass-Steagall, Erosion, and the 2008 Financial Crisis

In the wake of the 2008 Financial Crisis, there has been a lively debate among public figures and policy wonks about whether Gramm-Leach-Bliley’s “repeal” of Glass-Steagall exacerbated the financial crisis. For example, arguing for its guilt are House Majority leader Steny Hoyer (interviewed in Rowley and Harper 2009), and Ritholtz (2009); arguing for exoneration are former President Bill Clinton (interviewed in Bartiromo 2008), and Cowen (2008). Though my sympathies lie with the latter camp, plunging into these waters is beyond the scope of this chapter. Regardless, the account offered here provides reason to question the premises of some of the combatants. While there is no doubt that GLB was inimical to the original goals of Glass-Steagall, the 1999 law should be understood as much for its attempt to bridge the gap between law and reality as for its endorsement of expanded banking powers. GLB thereby helped to restore the rule of law to a policy arena that had become a bureaucratic free-for-all. Even those who believe the demise of Glass-Steagall made the financial crisis worse should acknowledge that the depression-era act’s erosion was a long process with many fathers, undertaken in concert with attempts to pursue competing regulatory goals. Holding GLB solely
responsible is rather nonsensical in light of this history. Indeed, Fisher (2001) creatively argues that on the eve of the passage of GLB, there was not a single Glass-Steagall prohibition that a company could not work around through the right maneuvers of corporate organization. Those who have attempted to indict GLB’s sponsors for Glass-Steagall’s demise should reconsider the identity of the culprits.

At the same time, it is easy to see how statutory erosion more generally could illuminate the onset of the financial crisis. Often discussed as a contributor to the crisis is the rise of a “shadow” banking sector in which all sorts of organizations increasingly utilized “exotic” financial instruments that escaped the confines of existing regulatory structures. This led to a complex web of obligations undermining the financial system’s stability, although few people recognized the situation’s precariousness before it was too late. In hindsight, such erosion of existing laws clearly demanded some sort of maintenance on the part of Congress to address. Figuring out exactly what kinds of maintenance would have made a difference is an extremely difficult task, though, as a strategy of “stringency on every front” might just have made escaping regulatory jurisdiction even more attractive to those meant to be included. There are powerful arguments that financial regulators, let alone legislators, are unlikely to possess sufficient information or dynamism to provide maintenance that is both timely and comprehensive enough to achieve systemic stability. If so, the best policies may be those that mitigate the ensuing harms when erosive innovation goes awry.

**Conclusion**

The preceding narrative shows why counteracting statutory erosion with maintenance is likely to be a difficult task, especially if bureaucrats’ priorities diverge, leading them to favor
erosive interpretations. In such circumstances, can any actions from the other branches be realistically expected to ensure the law’s continued efficacy?

Congress remains ever formally capable of providing maintenance, but its practical ability to do so may be limited, especially in policy areas where its involvement is sporadic. Where structural features of a law ensure the periodic involvement of legislators—either through requiring their reauthorization of a program or through a need to scrutinize its details for budgeting purposes—maintenance is far more likely to emerge from Congress, since the costs of paying close attention to the policy area will have to borne in any case. Maintenance will be most difficult, and thus require the most political will, when Congress has long been absent from the scene. In that case, as commercial bankers’ gradual accretion of powers in the context of Glass-Steagall showed, firms will acquire vested interests in the eroded regulatory regime and lobby against reforms that would reverse their gains. As these interests become more numerous and more vocal, cross-pressured legislators’ preferences may fail to coalesce enough to deliver statutory changes.

Legislation is hardly Congress’s only tool, however, and one might think that oversight remained a potent force to shape the ongoing interpretation of Glass-Steagall and the BHC. This essay’s narrative largely neglected this possibility because while congressional committee hearings certainly took place they seem to have had little effect. Probably the main reason is that the various arms of the banking bureaucracy are funded not through the standard congressional appropriations process, but through their own special assessments on banks. As an article in American Banker put it, bureaucrats might find themselves subject to a “tongue-lashing,” but if there was no real chance of a legislative follow-up, this could be shrugged off with relative equanimity. Representative Leach could thunder against the Comptroller’s abrogation of
Congress’s constitutional authority, but with every bill becoming tangled in the interest-group morass, this amounted to rhetoric (Seiberg 1998). Determining the extent to which legislative oversight in other policy areas has been able to achieve maintenance is a promising avenue for future research.

Where the legislature flounders in providing statutory maintenance, the judiciary may be best positioned to provide it by disciplining bureaucratic interpretations, but they are unlikely to be motivated to provide maintenance for its own sake. Political scientists often assume that judges have well-defined policy preferences and reach decisions which promote these. Whatever the plausibility of this hypothesis in the realm of politicized issues such as civil liberties, it is hard to discern any appreciable effect that individual judges’ specific policy preferences had in interpreting Glass-Steagall. Instead, judges’ overarching institutional concerns dictated their stances toward erosion and maintenance. Courts hope to define the scope of their own jurisdiction, defending their ability to weigh in but avoiding becoming arbiters of every conflict between regulators and regulated, a role that would swamp judicial capacity. Judges also want to create and support a coherent legal environment combining the capability to respond to changes with a measure of predictability. Flexibility and transparency are clearly in tension, and to the extent there is conflict among judges it seems to largely be over the optimal balance between these values. In many contexts, judges saw faithfully upholding the plain meaning of statutory language and Congress’s underlying intent as one important means of achieving stability, although figuring out exactly what that entails can be far from easy or uncontroversial. In other contexts, especially when congressional guidance was not forthcoming, coherence seemed more likely to be secured by deference to policy actors with greater expertise.
Judges recognize the difficulties inherent in their situation and try to make the best of things by looking to the capabilities of the other government actors involved. If Congress seems capable of staying involved in a policy area and dictating statutory maintenance or erosion as needed, judges are likely to adopt a textualist approach, narrowly interpreting statutory language. This is because Congress always retains the power to amend legislation, whereas this power will itself be undermined if courts allow themselves or bureaucrats to decide how statutes should operate even where there is no clear language to determine this.\(^\text{19}\) On the other hand, if legislators seem unwilling or unable to provide the necessary guidance, as they did in the 1980s and 1990s for Glass-Steagall policymaking, courts are likely to ratify bureaucratic solutions so long as they do not actively defy the terms of a statute—and perhaps even then, as long as the agency is duly respectful of legal niceties. Adapting the schematic table provided above again helps to organize the discussion (see Table 2).

The narrative recounted above features several combinations of these interpretive stances. In the early years, we saw Glass-Steagall given force when the executive and judicial branches both favored textualist maintenance. *ICI v. Camp* (1971) showed a bureaucratic textualist argument supporting erosion rejected by a Supreme Court willing to perform purposivist maintenance in service of what they believed were the act’s true purposes. In the early 1980s, the courts provided some textualist resistance to bureaucratic creativity in service of shifting ideas of the functions banking regulations should play, rejecting interpretations that lessened the force of Glass-Steagall as unfairly encroaching on Congress’s prerogative. After years of Congress failing to exercise that prerogative, though, the courts acceded to bureaucratic purposivist erosion.

\(^{19}\) This logic closely tracks the arguments made in the constitutional realm by Whittington (1999).
One of this paper’s most important conclusions is that when a statute comes under pressure, maintenance will only be effective if there is contemporary political will to re-commit to its original goals. If the most compelling counterargument against erosive interpretations a statute’s defenders can offer is that the plain meaning of the legal text has been given inadequate respect, that text will need to be crystal clear to stand against bureaucratic interpretations supported by expertise and resourceful government lawyers. “It’s against the law!” is, in the end, a rather feeble denunciation of government action if there is little independent belief in the law’s purposes. Given sufficient clarity in the statutory language at issue, judges with textualist proclivities may well rule in favor of such arguments, declaring that the law’s alleged failings are Congress’s to address. But defending a strict textualist reading of a statute becomes harder and harder as time goes by, more loopholes are discovered, and the law’s original goals and context become ever more remote (Strauss 2005). What executive branch interpreters make of the law becomes more important, with the content of the law providing less of a hard constraint on their choices.
Will a law’s erosion simply render law obsolete and irrelevant? The account provided here suggests otherwise: even as the efficacy of a law’s substance is eroded it will go on exerting formal requirements. As banking regulators admitted, without congressional intervention, respect for the rule of law made certain changes impossible. If erosion proceeds unaccompanied by statutory change, law can become a distortionary force, requiring costly compliance that no longer provides any of the substantive benefits originally envisioned.

The preceding discussion leaves many questions for further research. The discussion of purposivist interpretation has proceeded as if a statute contains a single underlying goal, or at least no conflicting goals. Often, however, the statute itself may articulate a function at an abstract level—such as serving the public good through ensuring a sound banking system—that seems to conflict with the goals animating its more specific commands. Employing the interest group theory of Stigler (1971), Macey (1984) argues that since most regulations are products of bargains between legislators and interest groups, judges should have little expectation that the actual function of laws will be consistent with the lofty, public-spirited goals offered as justification for the law. In the terms of this paper, judges motivated to effectuate public interest goals through their interpretations will end up undermining the true special interest purpose of the act. He concludes that judges should hard-headedly see the true special interest functions and respect them as Congress’s will, regardless of their ability to benefit the general public, although it seems just as easy to reach the opposite conclusion. In light of the importance of statutory maintenance to the healthy functioning of the regulatory state, it will perhaps be necessary to ponder a normative theory of how various institutional actors should orient themselves toward eroding statutes; certainly, this paper has shown that a bland injunction to “obey the law” won’t go far.
In general, it is worth asking how the nature of goals embedded in a statute affects judicial tendencies to provide statutory maintenance. For example, how does the judiciary cope with laws that explicitly aim to produce results that may be impossible to attain? Do these laws discourage maintenance by making it seem quixotic, or do judges take enacting legislators’ idealism to heart and provide as much support as the statute allows? What of statutes whose goals have been deliberately crafted to change with social beliefs, such that they encourage not only maintenance but active expansion without further legislative involvement?20

Finally, we lack a clear understanding of how legislators themselves think about statutory erosion and maintenance. Erosion should be a “known unknown” when Congress enacts legislation, a factor to be considered and planned for rather than left to future legislators likely to be distracted by other problems. But if we are all dead in the long-run, congressmen are out of office even faster than that. Studying when Congress chooses to take such concerns seriously, how they choose between legislative forms when they do so, and what information would help them more prudently anticipate and prevent statutory erosion, remains a promising avenue for future research.

---

20 See, for example, the argument of Sunstein (1998) in favor of allowing the FDA to assert its jurisdiction over tobacco after decades of rejecting such jurisdiction, because the agency could finally make a case that tobacco was indeed intended to provide users a pharmacological effect, which is the underlying statute’s intentionally open-ended criterion for inclusion under the regulatory regime. Such common law updating, Sunstein argues, is properly the work of agencies in the modern administrative state. This line of reasoning is largely reflected in Justice Breyer’s dissent in FDA v. Brown & Williamson Tobacco Corp., 529 U.S. 120 (2000), at 161. Calabresi (1982) makes an alternative argument, suggesting that the legal system ought to be adapted so as to return judges to their role as directors of common law evolution even in our current “age of statutes.”
Figure 1 – Bank Holding Company Corporate Organization (Each Unit’s Regulator Noted in Parentheses)

Parent Bank Holding Company (Federal Reserve)

Nonbank subsidiaries (Federal Reserve)

Out-of-State Bank Holding Company (Federal Reserve)

Nonbank subsidiaries (Federal Reserve)

National Bank (OCC)

Lead National Bank (OCC)

Nonbank subsidiaries (OCC)

State Member Bank (Federal Reserve)

State Non-Member Banks (FDIC and state)

Nonbank subsidiaries (OCC)

Lead National Bank (OCC)

Adapted from GAO (1995), Figure 1.

Cases Cited

Board of Governors v. Agnew, 329 U.S. 441 (1946)
Board of Governors v. ICI, 455 U.S. 551 (1981)
First Bancorp. v. Board of Governors, 728 F.2d 434 (10th Cir. 1984)
Dimension Financial Corporation v. Board of Governors, 744 F.2d 1402 (10th Cir. 1984)
ICI v. Conover, 593 F.Supp 846 (N.D.Cal. 1984)
ICI v. Clarke, 630 F.Supp 593 (D.Conn. 1986)
ICI v. Clarke, 789 F.2d 175 (2nd Cir. 1986)
ICI v. Conover, 790 F.2d 925 (D.C. Cir. 1986)
ICI v. Clarke, 793 F.2d 220 (9th Cir. 1986)
    Cert Denied, 479 U.S. 939 (1986)
SIA v. Board of Governors, 807 F.2d 1052 (D.C. Cir. 1986)
    Cert Denied, 483 U.S. 1005 (1987)
ICI v. FDIC, 815 F.2d 1540 (D.C. Cir. 1987)

Regulatory Materials Cited
Office of Comptroller of the Currency (OCC)
1957. Reprinted in Fed Banking L. Rep. (CCH) ¶ 96,262 at 81,357
    National Bank to Establish an Operating Subsidiary to Be Known as Security Pacific
    Discount Brokerage Services, Inc.” Quarterly Journal 1 (4): 40-44.
Board of Governors of the Federal Reserve (Fed.)
1983. “Statement by Michael Bradfield, General Counsel, Board of Governors of the Federal
    Reserve System, before the Subcommittee on Commerce, Consumer, and Monetary
    Affairs of the Committee on Government Operations, U.S. House of Representatives,
1984a. “Statement by Paul A. Volcker, Chairman, Board of Governors of the Federal Reserve
    System, before the Committee on Banking, Finance, and Urban Affairs, U.S. House of
    Representatives, June 12, 1984,” 70 Federal Reserve Bulletin 560, June.


United States General Accounting Office (GAO)


Reporting on Congress


Works Cited


